Date of acceptance Grade

Instructor

Guide on structure and layout of theses

Hannu Erkiö, Jaakko Kurhila, Lea Kutvonen, Matti Nykänen, Matti Mäkelä, Inkeri Verkamo

Helsinki September 12, 2007

UNIVERSITY OF HELSINKI
Department of Computer Science

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department

Tekijä — Författare — Author
Hannu Erkiö, Jaakko Kurhila, Lea Kutvonen, Matti Nykänen, Matti Mäkelä, Inkeri Verkamo

Työn nimi — Arbetets titel — Title

Guide on structure and layout of theses

Oppiaine — Läroämne — Subject

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
	September 12, 2007	10 pages + 1 appendices

Tiivistelmä — Referat — Abstract

This is a guide for the layout and structure of theses and essays at the Department of Computer Science at the University of Helsinki. It is thus applicable to the course Scientific Writing, the software engineering projects, seminars, and MSc theses.

ACM Computing Classification System (CCS):
A.1 [Introductory and Survey],
I.7.m [Document and text processing]

Avainsanat — Nyckelord — Keywords
layout, summary, list of references

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — övriga uppgifter — Additional information

Contents

1 Introduction

In all writing for publication, the writer's freedom of creation and expression are limited by a number of guidelines and specific regulations.

At best, a familiar set of regulations shared by reader and writer can create a kind of support network that allows the message to be relayed without distortion. It will be easier for readers to find the pertinent contents in a piece of writing if its layout and structure are the same as they are used to. This also applies to writers. When writers follow a set presentation model, they do not have to waste time on considerations that are secondary to the work itself, but they can concentrate on polishing the contents of the text. This means that it is a good idea to practice following the rules for the layout, though you may think you know how to select a better way to present your work.

This is a guide for the layout and structure of theses and essays at the Department of Computer Science at the University of Helsinki. It is thus applicable to the course Scientific Writing, the software engineering projects, seminars, and MSc theses. (This is an updated version of the previous guide written by the course lecturers [EMN01, ErM96, Erk94, Ver92].)

The LaTeX guide and LaTeX style that will soon be published on the department's web site can be used as support for this guide.

2 Structure

Let us start by looking at the sections expected to be in a scientific text. Keep in mind that the same expectations go for all kinds of technical writing.

2.1 Abstract or summary

The summary page contains the following sections: the bibliographical data of the work, an abstract, topic classification, and the key words. The bibliographical data consists of title, name of the author, place of publication, date of publication, and number of pages.

The abstract should be short, generally one paragraph (100 words maximum) explaining the main contents of the work: topic, methodology and results.

Topics are classified according to the ACM Computing Classification System (CCS). The lowest-level (third-level) descriptors should be used in this classification. In addition to a letter and number code, it is best to give the classification in plain text. Concepts that describe the main issues of the text should be selected as key words.

2.2 Introduction

The purpose of the introduction is to introduce the goals of the work in general terms. Describe topic, methodology and results (as in the abstract, but expand it).

The length of the introduction depends on the length of the whole work. A few pages of text does not need a separate introduction, since it is an expanded summary in itself. The introduction to a 10-page text can be 1–1.5 pages long. For a 50–70-page MSc thesis, a 2–4-page introduction seems reasonable.

The introduction should shortly describe the problem field of the whole work, the plot, and the conclusions, in general terms. After reading it, the reader may decide whether to go deeper into the topic by reading the whole text.

2.3 Topic chapters

The nature of the matter at hand determines how the topic chapters are disposed.

In order to guide the reader, it is a good idea to start each main chapter with a short paragraph on what the main topic of the chapter is and how it progresses from one sub-chapter to the next.

Texts with only one sub-chapter, or with more than two chapter levels (main and sub-chapters) are a sign of a problem with the disposition of the text. There may be justifiable reasons to use three-level headings in some technical documents, but they are an exception to the rule.

2.4 Reference usage

So-called mnemonic references are used for referring to sources; they are constructed as described in the section on the list of references. The page numbers should be added to the reference if it would be too laborious for the reader to find the reference in the source without them.

References are always placed inside sentences. This means that e.g. a separate reference at the end of a paragraph would be inappropriate.

The structure of the text must clearly show what the reference relates to. At the same time, it shows how long a piece of the text that the reference relates to.

2.5 Conclusion

At its simplest, a conclusion may be a short revision of the main points of the text. A more valuable conclusion contains comments on e.g. the value of results, how the work relates to its environment, or future visions. This kind of evaluations should be well-grounded in fact, though, or the conclusion might inadvertently seem comical.

2.6 Creating the list of references

The following guidelines should be followed when creating lists of reference for the assignments during the course Scientific Writing.

The guidelines are backed by two main goals: to make it as easy as possible to find the referenced source, and to show what kind of evaluation process the referenced work has undergone. For these reasons

- the reference notes should always be so exact that the source can be recognized and found in catalogues and libraries
- different types of sources (monographs, conferences, journals) have to be easy to distinguish from each other, and
- the different parts of the list must conform to each other, especially for each source type.

The sources are listed alphabetically according to the author's name, and works by the same author (group of authors) in the order in which they have been published. If some publication does not have an individual writer, it is alphabetized according to its name. A mnemonic abbreviation consisting of three letters and two numbers is created to identify the source. The letters are the first letters of the author's or authors' surname(s), and the numbers are the last two digits of the year. If there a re several publications from the same author and the same year, they are identified with a lower-case letter placed after the year.

The following information should be given on each source:

- Name(s) of author(s) (surname, initial letters of first names) in the original order; if there are more than three authors, we can write the name of the first author and *et al.* instead of the other names.
- the name of the publication or article in its original form
- place of publication:
  - of monographs: publisher, place of publication (can be omitted if the publisher is well known), year
  - of journal articles: name of journal, volume, issue, year and month (in parenthesis)
  - of articles from article collections (such as conference publications):
    - name of collection, editor, publisher, place of publication and year *or*
    - conference name, coordinator, place and time,
  - of a report: series, report number, place, publisher and year
  - of a web source: URL, validity date, possibly the date when referenced in square brackets
- page numbers, if the source is an article or constitutes a chapter in a compilation.

Put a full stop after the name of a publication or article, as well as after the bibliographical data of each reference. Separate the other pieces of information with a comma. As is normal in Finnish, only the first letter of the first word in the heading is capitalized, but in the titles of conferences and compilations, each major word is capitalized (not articles or prepositions). See the appended example for a model. For the sake of clarity, it is best to write *In the work* before the name of a compilation, except in the case of conference publications where the name starts with the abbreviation *Proc.* (for Proceedings). In such cases, no complement is necessary. You can see the difference by comparing the layouts of the references "[DaT90]" and "[Gan89]" .

When using web sources, you should keep in mind that the threshold for publication on the web is non-existent. It is better to concentrate on the publications of well-known scientific publishers and the technical standards for which the web is the

only publication channel. If the same publication is available on paper, refer to that primarily and add the URL as a complement.

The list of references gives an example of a text that has been published through many channels [AQM97, Die99]and another example that shows a standard that has been disseminated only through the web [BPS98].

There is a separate guideline, [Kil00], on using and referring to web sources.

The list of references for a text should list exactly the sources that the text refers to. The list of references for this text is an example of how to present sources, which is why it contains ”extra” sources.

3 Layout

This chapter discusses the main issues of the technical presentation of a text.

3.1 Disposition of text sections

Each text should include a separate cover page, as in this model. The second page contains an abstract, followed by a table of contents (on one or more pages), and then the main text. Pagination starts on the first page of the main text (with the Arabic numeral 1). (Rigorous writers leave out the number from the first page.) All the (numbered) headings and their page numbers should be written into the table of contents. Many word-processing systems create the table of contents automatically so that the writer does not have to worry about updating page numbers as the work progresses. If writers wish to, they can paginate the table of contents and previous pages separately (with Roman numerals), e.g. like this model.

After the main text, but as part of the text body, comes the list of references; its heading is not numbered. Any possible appendices are added after the list of references, with headings and internal page numbers.

If you want to make a coherent list of figures, algorithms and tables, this list should be placed i mmediately after the table of contents. The value of such lists is a matter of opinion, so there is no need to create one — especially if your word-processing software does not support it — unless your instructor explicitly asks for it.

If for some special reason you want to add an alphabetical index top your work, it should be placed after the list of references and before appendices. An index should

be noted in the table of contents, as s hould the list of references (unnumbered chapter). Writers who want to create an index should use the automatic support their word-processing system offes.

3.2 General text layout

Print on one side of the paper only. Select sparse line spacing (1.5–2) so that your instructor can write comments between the lines. You can use even sparser line spacing in your own versions, if necessary. Both margins should also be fairly broad (ca. 3 cm): the left margin is needed for binding the text and the right one for the instructor's comments. Leave enough space (2–3 cm) at the top and bottom, as well.

The most effective way to distinguish features like chapters, figures, etc, is to have enough space in your text. Separate paragraphs from each other with one and chapters with a few empty lines. If a new chapter is about to start at the bottom of a page (with only one or two lines of text), it is better to start it on the next page. However, it is not necessary to start every chapter on a new page, especially not with a short text; if the text contains many pages that are nearly empty, readers might suspect that the writer has tried to make it look longer than it is.

Any empty space may be utilized for displaying figures and tables. Especially if a text is written with the same type of text throughout, empty lines are necessary for separating e.g. text from tables. Empty space is cheap, but adds to the clarity and readability of a text.

3.3 Figures and tables

All figures and tables should be placed as near the (first) place in the text that refers to them, but not before this reference. The text should also explain what the writer wants to illustrate with the figure. Figures can be interpreted in many different ways, so the reader needs guiding.

Figures should never be placed immediately under the heading of a chapter, but chapters should always start with text. Figures should not be placed in the middle of a paragraph (much less a sentence), except if the figure is placed at the top or bottom of a page and it is clear that the paragraph continues.

Figures do not always have to be placed immediately after the paragraph that refers to them. If there is not enough space for a figure on the same page as the paragraph

that refers to it, the rest of the page should not remain empty, Though the figure is inserted on the following page. However, the figure should never be more than one page away from the reference.

The image 1 shows how to present a figure. You must pay attention to the visibility of figure parts and text, to the numbering of figures, and captions.

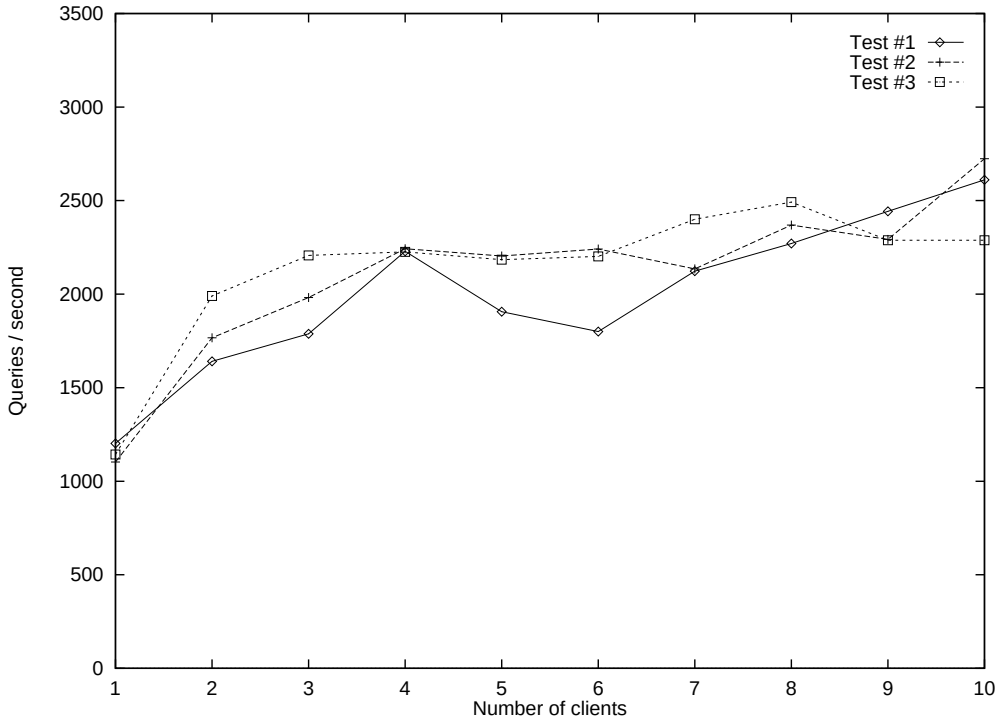

Figure 1: Figure elements.

We must also pay attention to the size of figures. Any annotations must be clear and easy to read When presenting performance curves, for example, the axels have to be named, the scale notated, and the units clearly presented. If you present many things with similar figures, you should use the same scale for easy comparison.

The caption of a figure should be written underneath it, and it is preferable that it be short and to the point than too explicit. The same goes for table captions.

Figures and tables should be numbered progressively. In long texts, two-level numbering should be used (e.g. Figure 3.1) according to the chapter number, but in shorter texts, one-level numbering is good enough.

You should pay attention to presenting figure and table captions consistently, as

well as to punctuation marks. It is natural to put a full stop after a caption, since they are most often full sentences.

(The style for figure and table captions vary according to publisher and publication. The recommendations at the Department of Computer Science also seems to vary with respect to where the caption should be placed.)

3.4 Headings

You can use a different font in headings than in the rest of the text, or underlining, larger fonts, or other methods of emphasis, but generally it is preferred that you only use one of these methods, because if there are very many font types and sizes, the layout looks messy. The format of headings must be consistent throughout the text. You should not use any unnumbered "extra" headings.

3.5 Using this model

You can use this text as a model for the layout of your own assignment. The font types and sizes, line spacing, etc vary according to word-processing system, so small deviations from the rule are acceptable.

The directive number of pages for written assignments given during the lectures for the Scientific-writing course and in this guide are applicable to assignments with a layout like this guide's (font size 12 points). The average line in this text should consist of about 80 characters and one page about 30 lines. The number of pages includes the text itself and the list of references (the part paginated with Arabic numbers), not the cover page, summary, or table of contents.

4 Conclusion

This text is a checklist for some of the rules governing written presentations, which you should keep in mind when writing exercises and theses.

References

AQM97 Abiteboul, S., Quass, D., McHugh, J., Widom, J. and Wiener, J., The lorel query language for semistructured data. *International Journal on Digital Libraries*, 1,1(1997), pages 68–88. [Myös `http://link.springer.de/link/service/journals/00799/bibs/7001001/70010068.htm`, 18.1.2000].

BPS98 Bray, T., Paoli, J. and Sperberg-McQueen, C., Extensible Markup Language (XML) 1.0. W3C Recommendation 10-February-1998. `http://www.w3.org/TR/1998/REC-xml-19980210`. [18.1.2000]

CDM98 Crowder, H., Dembo, R. and Mulvey, J., On reporting computational experiments with mathematical software. *ACM Transactions on Mathematical Software*, 5,2(1979), pages 193–203.

Die99 Dietinger, T. et al., Dynamic Background Libraries - New Developments in Distance Education Using HIKS (Hierarchical Interactive Knowledge System). *Journal of Universal Computer Science*, 5,1(1999). [Myös `http://www.iicm.edu/jucs_5_1/dynamic_background_libraries_new`, 18.1.2000].

DaT90 Dan, A. and Towsley, D., An approximate analysis of the LRU and FIFO buffer replacement schemes. *Proc. ACM Conf. Measurement and Modeling of Computer Systems*, Boulder, Colorado, USA, May 1990, pages 143–152.

ErM96 Erkiö, H. and Mäkelä, M., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelytieteen laitos, Helsinki, 1996.

EMN01 Erkiö, H., Mäkelä, M., Nykänen, M. and Verkamo, I., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelyopin laitos, Helsinki, 2001.

Erk94 Erkiö, H., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelyopin laitos, Helsinki, 1994.

FHS89 Fogelberg, P., Herranen, M. and Sinikara, K., *Tuumasta toimeen, tutkielman tekijän opas.* Yliopistopaino, Helsinki, 1989.

Gan89 Gannon, D. et al., Programming environments for parallel algorithms. In *Parallel and Distributed Algorithms*, Cosnard, M. et al., editors, North-Holland, 1989, pages 101–108.

Gri87 Grimm, S. S., *How to write computer documentation for users.* Van Nostrand Reinhold Co., New York, 1987.

HaP82 Harkins, C. and Plung, D. L., editors, *A guide for writing better technical papers.* IEEE Press, 1982.

Jul81 Julkaisusarjoja ja opinnäytteiden tiivistelmiä koskevat ohjeet ja suositukset. Helsingin yliopiston kirjastolaitoksen julkaisu A 3, Helsinki, 1981.

Kil00 Kilpeläinen, P., WWW-lähteisiin viittaaminen tutkielmatekstissä, 2000. `http://www.cs.helsinki.fi/u/kilpelai/TiKi/wwwlahteet.html`. [19.1.2000]

Smi78b Smith, A. J., Sequential program prefetching in memory hierarchies. *Computer*, 11,11(1978), pages 7–21.

Smi78a Smith, A. J., Sequentiality and prefetching in database systems. *ACM Transactions on Database Systems*, 3,3(1978), pages 223–247.

Ver92 Verkamo, A. I., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelyopin laitos, Helsinki, 1992.

Appendix 1. Model ABC

The appendices here are just models of the table of contents and the presentation. Each appendix usually starts on its own page, with the name and number of the appendix at the top. Each appendix is paginated separately.

In addition to complementing the main document, each appendix is also its own, independent entity. This means that an appendix cannot be just an image or a piece of programming, but the appendix must explain its contents and meaning.