582519 Scientific Writing for MSc in Computer Science: Writing process

Lecture 2, 10.9.2013
Tiina Niklander
Last week’s task: How did it go?

Introduction to the Use of Computers?
C programming?
Other courses?
Small group meetings?

What did you do last week for your paper?
Course material

Read one book about scientific (or academic writing):
• Zobel: Writing for computer science
• Murray: How to write a thesis

Read general style guides:
• Strunk: The elements of style (1st ed 1918, 5th ed 2009)

Any general books about writing in English:
• Grammar, punctuation,
What is scientific writing?

- Describing new scientific ideas and results for other scientists and the science community
  - New ideas or combining old ideas in a new way
- Evaluation of the presented ideas/results
  - Correctness
  - Relationship to former ideas and results
- Scientific writing is an important part of the process of science, i.e., the process of accumulating reliable knowledge
What is scientific writing? (2)

- New results (new contribution) build on existing knowledge
- Scientific results are accepted when they have been independently reviewed and published
- Different types of scientific publications with their own characteristics

- Scientific writing itself is also a process
The process of science

- Idea or research question
- Hypothesis or model for the solution
- Testing the hypothesis/model
- Representing and evaluation of the results
- Publishing the results after independent reviewing

- Good writing is a crucial part of this process!
## Bloom’s Taxonomy Action Verbs
by Clemson university

<table>
<thead>
<tr>
<th>Level</th>
<th>Bloom’s definition</th>
<th>Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Remember previously learned information</td>
<td>Memorize, <strong>list</strong>, define, identify, repeat, order</td>
</tr>
<tr>
<td>Comprehension</td>
<td>Demonstrate an understanding of the facts</td>
<td>Classify, <strong>paraphrase</strong>, explain, extend, select</td>
</tr>
<tr>
<td>Application</td>
<td>Apply knowledge to actual situations</td>
<td>Change, employ, use, modify, <strong>relate</strong>, discover</td>
</tr>
<tr>
<td>Analysis</td>
<td>Break down objects or ideas into simpler parts and find evidence to support gener.</td>
<td>Appraise, criticize, model, examine, <strong>diagram</strong></td>
</tr>
<tr>
<td>Synthesis</td>
<td>Compile component ideas into a new whole or propose alternative solutions</td>
<td>Arrange, create, plan, collect, develop, <strong>write</strong></td>
</tr>
<tr>
<td>Evaluation</td>
<td>Make and defend judgments based on internal evidence or external criteria</td>
<td><strong>Argue</strong>, assess, choose, compare, evaluate</td>
</tr>
</tbody>
</table>

http://www.clemson.edu/assessment/assessmentpractices/referencematerials/
<table>
<thead>
<tr>
<th>LEVEL</th>
<th>FOCUS</th>
<th>GOAL</th>
<th>VERBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INFORMATION</td>
<td>Identification and repeating - able to repeat accurately</td>
<td>Know that you know</td>
</tr>
<tr>
<td>2</td>
<td>UNDERSTANDING</td>
<td>Reach the impact and meaning of information</td>
<td>Show that you understand</td>
</tr>
<tr>
<td>3</td>
<td>APPLICATION</td>
<td>Use information - apply knowledge to new situations</td>
<td>Show: use the information</td>
</tr>
<tr>
<td>4</td>
<td>ANALYSIS</td>
<td>Make conclusions f.e. show relationships between elements and the whole</td>
<td>Show: find the essential elements in the information</td>
</tr>
<tr>
<td>5</td>
<td>SYNTESIS</td>
<td>Be creative and original - compose elements and ideas in new ways</td>
<td>Show: create and formulate large wholes</td>
</tr>
<tr>
<td>6</td>
<td>EVALUATION</td>
<td>Create criteria in order to evaluate the information. Use the criteria.</td>
<td>Show: evaluate ideas, information, methods and solutions</td>
</tr>
</tbody>
</table>
Contents of a scientific article

- The position of the new idea with respect to the former scientific knowledge
- Clear and formal statement of the new idea
  - Often as a hypothesis or a model
- Description of the novelty of the idea, or the contribution of the article
- Justification of the theory by proof or experiments

- Note: details of the research process leading to the article do not usually belong to the article!
Structure of a scientific article

- Title and author
- Abstract
- Introduction
- Body of the article (including related research/work)
- Summary/conclusions
- (Acknowledgements)
- Bibliography
- Appendices
Types of publications

- Course books
- Other kinds of text books
- Standards
- Journal articles
- Conference articles
- Theses
- Technical reports
- Manuscript
- Web documents/publications
Nature of publications

- **Primary sources**
  - Original, reliable information
  - Articles in scientific journals, conference and workshop books, research reports, theses
  - Patents, laws and directives, standards

- **Secondary sources**
  - Useful material for restructuring and analysis of existing information, not original and new results
  - Course books, article collections, surveys, newsletters, dictionaries, etc.
Process of writing scientific text

- Idea or interesting topic
- Finding and evaluating relevant source material
- Reading material
- Identifying essential issues
- Restructuring them logically
- Writing them down using proper presentation techniques
- **Iterative process**: text must be re-written several times!
Characteristics of scientific text

- Content usually technical
- Should be based on facts
  - Writer's interests and opinions can be seen in the choice of the topic, not in the text as such (student paper’s normally contain justified opinions also)
- Based on former theories and research results
- Motivating on why the problem considered is important
  - Not marketing of the ideas
- Arguments and conclusions
- Verifiability, reliability, and repeatability of the results
Characteristics of scientific text (2)

- Clear and logical structure
- Not a direct copy from anybody else's text
- Source material is analysed and restructured
- Based on peer-reviewed research material
  - Journal, conference and workshop articles in computer science are typically peer-reviewed
  - Peer-reviewers are researchers that are experts in the topic in question
  - Writers do not know who the reviewers are
Characteristics of scientific text (3)

- Text is suitable for its target group
  - How are the readers?
  - How are they going to use the text and the information given in it?

- Clearness of the text
  - The reader must understand the text in a same way as the writer

- Reflects writer's deep understanding of the topic!
Target groups of scientific text

- Readers that have scientific background
- Other researchers in the same area
- Whole scientific community
- General public
- Some basic knowledge of the topic is usually required

- In this course and in the seminars: other MSc degree students
How do you learn scientific writing?

- Following the topic area and reading relevant articles
- Writing yourself
- Searching for feedback from others
  - Peer students
  - Teachers
  - ...
- Iterative process!
Important things to remember

- It is important that the reader understands what you have written
  - The audience you should write to are the other students in this course!
- All the texts must be written by you
  - Not a copy of a text written by someone else
  - Not a direct copy from the reference material
- Remember always to check the correctness of the language!
Important things in your paper

- Department’s layout for thesis and reports
- Structure of your text
- List of references
How to start: Collecting joint ideas

- -- to be added in the class --
Writing process
Writing process

- (Scientific) writing is a process

- Different phases
  - Thinking and planning
  - Outlining
  - Writing
  - Revising
  - Finalising

- Iteration between phases sometimes necessary
Thinking and planning

- Getting started
  - Defining purpose and scope
  - Identifying target group

- Gathering ideas
  - Finding relevant background information
  - Identifying things to be included

- Taking the requirements of the intended publication forum into account
Getting started

- Define the purpose and scope of your text
  - What are the problems considered? Why are they interesting?
  - What perspective is taken?
  - What are the objectives of the text?
  - Which purpose the text is written to?

- Who are the readers? (target group)
  - What the readers want to know?
  - What information should be given to them?
  - In this course: your fellow students
Gathering ideas

- Finding background information
  - Information retrieval of scientific text
  - Reading the relevant material
  - Defining the main references to be used
  - What kind of related work should be considered?

- Identifying things (topics) to be included and finding their relationships
  - Key words, key phrases
  - Relevant concepts and definitions
  - The most important results to be presented
Outlining

- **Starting point**: What is text's scope and purpose?

- What does the reader need to learn or know about the topic?

- Starting from the common knowledge and proceeding to the new ideas and results

- **Logical structure and order** of the text is essential!
Outlining (2)

- Helps authors to
  - organise their thoughts
  - evaluate relevance of different topics and their representation
  - remember the relationships between topics
- Should **support reading** and reader's understanding
- Describes
  - structure of the text
  - logical presentation and reading order
- Should still support several types of reading
  - browsing, specific information searches, learning, ...
Outlining (3)

- Should lead to a logical, clear story
  - In a concise form in the list of contents
  - Clarified in the introduction, especially if there is something special in it

- Questions to answer:
  - What topics are considered?
  - In which order they are told?
  - What is the importance and length of each topic?
Outlining (4)

- What kind of parts are needed?
  - Which chapters?
    - Only in longer texts; seldom in scientific articles
  - Which sections?
  - Which subsections?
  - Some other parts?

- Finding
  - a good title for the whole text
  - headings for chapters, sections and subsections
Outlining (5)

- Some publication forums may have strict rules on the outline
  - For example: Introduction, Methods, Results, Discussion

- Can cause problems in explaining complex topics in phases
  - For example a comparison of two methods => Introduction, Background, Methods, Results, Discussion, Methods, Results, Discussion

- Not typical in computer science
Different types of outlines

- Chaining outline
  - Presentation of the problem
  - Related work, earlier solutions and their flaws
  - New solution
  - Results and their evaluation

- Specificity-based outline
  - First general explanation/description, then more specific ones
  - For example for describing a system consisting of several components
Different types of outlines (2)

- Example-based outline
  - Idea or results explained first with help of a typical case or situation
  - Generalisation of ideas/results and describing them more formally

- Complexity-based outline
  - First presentation of a simple case
  - Then description of a more complicated case (generalisation, extension)
A title of an article/thesis/report must be informative and concise
  - Too general terms and titles should be avoided
  - Every term should be necessary

Must be attractive
- Not too complicated and filled with words
- Not too short either

Preciseness is more important than conciseness and attractiveness!
Titles (2)

- Examples:
  - Too complicated:
    - An Investigation of the Effectiveness of Extensions to Standard Ranking Techniques for Large Text Collections
  - Better:
    - Extensions to Ranking Techniques for Large Text Collections
  - Too general:
    - Huffman Coding for Databases
  - Better:
    - Limited-Memory Huffman Coding for Databases of Textual and Numeric Data
Chapter and section headings

- Should reflect the structure of the work
  - For example
    - 4. List and trees
      - 4.1. Lists
      - 4.2. Trees

- Not complete sentences
  - Example:
    - Not: Replication of Data Leads to Reduction in Network Traffic
    - But: Replicating Data to Reduce Network Traffic
Chapter and section headings (2)

- Not too lively
- Avoid questions or abbreviations

- Headings at the same level should
  - be comparable in their contents and structure
  - have a clear connection to the balanced outline

- Third-level headings, i.e. subsubsections, seldom needed
  - Usually the need of them indicates problems in the outline
Chapter and section headings (3)

- Paragraph titles should be avoided
  - If needed, should be part of the paragraph

- Numbering of headings depends on the publication forum
  - Unnumbered headings must be distinguished by a specific font, style or font size
  - At our department numbering of headings is required
Paragraphs

- Building blocks of chapter, sections and subsections
- Should not be too long
  - Logical flow of the text becomes difficult to follow
- Short paragraphs easier to read and they make communication more efficient
  - No paragraphs consisting of just one sentence!
- A paragraph for each aspect of the topic
Scientific text should be impartial, accurate and objective
- Arguments must be based on evidence
- Statements should be supported by examples
- Sources of information and ideas must be indicated
- Use enough words to make your meaning clear

Started by writing a draft of the text
- Flow of ideas
- A short text can be drafted completely
First draft

- Freely written
  - Concentrate on presenting ideas in a logical way
- Raw text
  - Style, layout and punctuation can be corrected later
  - Exception: mathematical and formal issues as precisely as possible from the very beginning
- Must be edited and revised carefully and thoroughly
  - Several times
  - Difficult things more times than easier
How to proceed with writing

- Different approaches
  - Write the introduction first
  - Start from the body of the text
  => use the method that is the best for you

- Write something even if it is hard
- Start with easier things
- If everything else is difficult, fix the technical details (list of references, etc.)
Revising

- After the first draft is ready/complete

- **Aim:** ensure that thoughts created in the mind of the reader(s) are the same as the thoughts of the writer(s)

- Checking
  - the order of presented ideas
  - the use of words and terminology
  - style, layout and punctuation
Revising (2)

- No statement should be introduced abruptly and without warning

- Relationships between parts at a same level should be clear
  - Each section should be related to the preceding and the next one
  - Similarly with paragraphs in a section and sentences in a paragraph
Revising (3)

- Paragraphs and sentences should be in a logical and effective order

- Balance is important
  - Parts must be balanced in themselves, and in the relation to one another
  - Holds for sections, subsections, paragraph, and even sentences
  - For example, no sections with just one subsection!
Revising (4)

- Important and difficult parts typically re-written several times

- After a revision, put the text a side for a moment
  - Avoid blindness to your own text!

- Ask someone to read your text and give comments!
  - Experts versus non-experts
  - The function of criticism and feedback is to improve your writing
When all the parts of the text are written at least once
Check that objectives, motivation and restrictions are in line
Evaluate issues that the readers or reviewers might criticise or argue against
Check technical details
If published, the critics of the reviewers should be taken into account
Changing publication forum may require bigger changes
Structure of a thesis, report or seminar paper
Structure of a thesis, report or seminar work

- Title page
- Abstract
- Table of contents
- Introduction
- Body of the work
- Conclusions
- List of references
- (Appendices)
Abstract

- Helps the reader to decide whether to read the whole text or not
- Should be short, but clear, informative and concise
- Details or description of the outline not part of the abstract
- Final version is written after all the other parts of the text are ready
- Aimed to large audience
  - Readers are not necessarily experts on the topic
- No references to other articles
Introduction

- Introduction should be the easiest part of the text to read and not too long
- Must tell what are the new ideas and the main results presented in the article/report/thesis
- Must show that that the whole text is worth reading
- Can be written first, but then typically must be revised
Introduction (2)

- Not too technical, but specific and informative enough
  - Not technical terminology
  - Not very detailed definitions of terms/concepts
  - Not too much mathematics and formulas

- Must contain motivation
- No deep literature analysis
- Importance of the results, not conclusions
- (References to relevant work)
Introduction (3)

- Typical contents and structure
  - Description of the topic and its context
  - (Related work)
  - Description of the problem considered
  - Summary of the proposed solution
  - Evaluation of the solution
    - Application areas
    - Consequences
  - Brief description of the outline of the work
- In surveys: why the certain approaches were chosen
Literature overview

- Related work, survey
- Connections to earlier research on relevant topics
- As important as the description of the contribution of the text
- Location in the text
  - In the beginning of the article (description of the context, a part of introduction)
  - If large, a separate section is required
    - As a part of the body of the text
    - After the body, where a comparison of old and new solutions is possible
  - In different sections when it is appropriate
Definitions

- Terminology, variables, abbreviations and acronyms must be defined or explained **the first time** they appear in the text.
- Consistent emphasising
  - Different style of letters: italics, boldfacing, …
  - Only the first occurrence
- Sometimes several explanations can be good
- Definitions are given when needed
  - Usually a separate section “Definitions” is not needed/good
  - Every defined term should be necessary
Results and their analysis

- Traditional order of presentation:
  - Description of all results
  - Analysis of the results

- Drawback: the reader might not be able to follow what happens

- More reasonable order of presentation:
  - Analysis is combined with the description of results and how they are obtained

- Description of a particular result should usually start with a brief summary on the main observations
Conclusions

- Brief repetition of the main ideas, results and conclusions as well as their meaning
- Restrictions of the work can be repeated
- No new ideas or conclusions that are not presented in the body
- Can be stated
  - Unsolved problems
  - Which points or perspectives were omitted
  - Which variations should be considered/researched further
Divided authorships

- All the authors have some kind of contribution to the contents of the article
- Brainstorming and developing ideas
- Even writing together
  - Each author write a certain part of the text (different styles, non-coherent style)
  - One or two authors write the draft, and other revise it in turns