



TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

Technical Writing and Presentation

Writing scientific text in computer science

SOICT - 2020

Scientific Writing for Computer Science

1. How to write scientific texts in computer science?
2. How to write in English?
3. How to write a technical report / thesis?

Contents

- General style
- References and citations
- Acknowledgements
- Formatting
- Use of tables, figures, and similar elements.
- Mathematics

General style

- What is style?
- Scientific writing style

Genre (style)

- The manner of expression is the writing style: how well you communicate with likely readers.
- Scientific writing must be plain and straightforward but not dull. It can have style.
- Good style makes your writing easy to understand. It can persuade readers that work is of value.

Scientific writing style

- Fundamental aims of science writing: to be exact, clear, and compact.
- Compact is usually clear!
- Other desirable properties: smooth and objective

- Word choice
- Avoid vague expressions which are typical for the spoken language
- Make clear what the pronouns refer to
- Avoid ambiguity
- Avoid illogical comparisons
- Correct grammar

Avoid ambiguity

- Example

✗ An inverted list for a given term is a sequence of pairs, where the first element in each pair is a document identifier and the second is the frequency of the term in the document to which the identifier corresponds.

✓ An inverted list for a term t is a sequence of pairs of the form (d, f) where each d is a document identifier and f is the frequency of t in d

- Many terms have well-defined mathematical meanings and are confusing if used in another way.
- For example,
 - Formula, expression, equation
 - Average, mean
 - Subset, proper subset, strict subset.

Use Words Efficiently

- Never use two words when one word will do.

Example 1: The relationship between the nature of salt water to fresh water in the **Edgartown Great Pond** that fluctuates often is extremely important to everyone including scientists, residents, and environmentalists on Martha's Vineyard.

Example 2: The fluctuating salinity of **EGP** concerns many environmentalists, scientists, and residents.

Clear

- Use illustrative titles which describe the essential in a chapter or a section.
- Write a brief introductory paragraph in the beginning of each chapter or section with subsections.
- Divide the text logically into sentences and paragraphs.
- Place the adjective or the adverb as close as possible to the word it modifies.
- Avoid **scientific jargon** = continuous use of technical vocabulary when it is not relevant.
- Write numbers as digits when they refer to sizes or exact measurements.
- Use punctuation to support meaning.

Avoid jargon

- Jargon: a vocabulary particular to a place of work (abbreviations, slang)
- Audience familiarity with the topic determines appropriate use of jargon
 - ✗ For the first year, the links with SDPC and the HAC were not connected, and all required OCS input data were artificially loaded. Thus CATCH22 and MERWIN were not available.
 - ✓ Because some of the links in the computer system were not connected the first year, we could not run all the software codes.

Define the Unfamiliar

- If you must abbreviate, define the term in its first occurrence, and ***put abbreviations in parentheses***

Example: Edgartown Great Pond (EGP) is a vital body of water. Unfortunately, due to an unpredictable influx of saltwater, the delicate ecosystem is in danger of destabilizing.

- ***Italicize first occurrence*** of unfamiliar terms and define them right away

Example *Retina* is a light-sensitive tissue, found at the back of the eye, that converts light impulses to nerve impulses.

Compact

- Say only what needs to be said!
- Short words and short sentences are always easier to comprehend
- Weed out too detailed descriptions.
- Don't describe irrelevant or trivial observations
- Avoid wordiness, Use no more words than are necessary. Redundant words and phrases should be omitted.
- Avoid too long sentences and paragraphs

Smooth

- Verbs: Stay within the chosen tense! No unnecessary shifts in verb tense within – the same paragraph – in adjacent paragraphs
- Use verbs rather than their noun equivalents
- Prefer active to passive voice
- Avoid long noun strings!
- Each pronoun should agree with the referant in number and gender.
- Transitional words help to maintain the flow of thought

Smooth

- Notice: some transitional words (while, since) can be used in several meanings → limit their use to their temporal meaning!
- Use abbreviations sparingly, especially the abbreviations which you define yourself for technical terms.
- Do not use emphasis (italics) when it is not needed. Use syntax to provide emphasis.
- Metaphors can sometimes help to simplify complex ideas. However, – Avoid overuse metaphors and cliches

Active voice

- Use direct statements and expressions involving “we” or “I”—that is, the active voice—to make reading more pleasant and to help distinguish new results from old.

Example: active voice vs passive voice

✗ The results show that, for the given data, less memory is likely to be required by the new structure, depending on the magnitude of the numbers to be stored and the access pattern.

✓ The results show that less memory was required by the new structure. Whether this result holds for other data sets will depend on the magnitude of the numbers and the access pattern, but we expect that the new structure will usually require less memory than the old.

References and Citations

- References and citations
 - explain the relationship of your new work to existing works.
 - how your work was built on previous knowledge and how it differs from contributions in other paper
- References must be in standardized format

Purposes of references

- Demonstrate that your work is new
- Demonstrate your knowledge of the research area
- Help the reader to judge whether your statements are reliable.
- Pointers to background reading.

An appropriate reference is

- Relevant
- up-to-date
- reasonably accessible
- Refer to the original paper
- well-written material
- book, conference paper, journal article in preference to a workshop paper
- workshop paper in preference to a manuscript
- formally published documents rather than Web pages
- Avoid reference to private communications and information provided in seminars or talks
- Don't cite to support common knowledge.

Examples

- Refer to a result in an inaccessible paper, do not refer to it directly:
According to Kelly (1959; as quoted by Dawson 1981), stable graphs are closed
- When you use results from other papers, be sure to show the relationship to your own work:
 - Marsden [16] has used an approach in which...
Other work (Marsden 1991) has used an approach in which...

Citations

- Citations should be punctuated
- Never treat a bracketed expression, whether a citation or otherwise, **as a word**.
 - ✗ In [2] such cases are shown to be rare.
 - ✓ Such cases have been shown to be rare [2].

Citations

- The cite should be close to the material it relates to

✗ The original algorithm has asymptotic cost $O(n^2)$ but low memory usage, so it is not entirely superseded by Ahlberg's approach, which although of cost $O(n \log n)$ requires a large in-memory array (Ahlberg 1996; Keele 1989)

✓ The original algorithm has asymptotic cost $O(n_2)$ but low memory usage (Keele 1989). Thus it is not entirely superseded by Ahlberg's approach (Ahlberg 1996), which, although of cost $O(n \log n)$, requires a large in-memory array

Acknowledgements

- Should thank everyone who made a contribution, whether advice, proofreading, coding, or whatever
- Two common forms of acknowledgement
 - list the people who have helped
 - explain each person's contribution
- Write “I am grateful to” , “I thank” or “Thanks to” instead of “I would like to thank” or “I wish to thank”

Formatting

- Fonts
- Indentation
- Layout
- Headings
- Numbering

Fonts

- Standard choices: Times New Roman, Cambria.
- Three fonts for computing and mathematical writing : plain, italic, bold
- Underlining is used for emphasis.
-

Indentation

- Indentation is used to indicate the start of a new paragraph
- Use right-justified as well as left-justified
- Pages should be numbered

Layout

- The appearance of a report is no less important than its content.
- An attractive, clearly organised report stands a better chance of being read.
- Use a standard, 12pt, font, such as Times New Roman, for the main text.
- Use different font sizes, bold, italic and underline where appropriate but not to excess.
- Too many changes of type style can look very fussy.

Headings

- Use heading and sub-headings to break up the text and to guide the reader.
- They should be based on the logical sequence which you identified at the planning stage but
- With enough sub-headings to break up the material into manageable chunks.

Numbering

- The use of numbering and type size and style can clarify the structure as follows;

3 Methods of harnessing wave energy

3.1 Shore-based systems

3.2 Deep-water systems

3.2.1 *"Duck"* devices

3.2.2 Rafts

Use of tables, figures, and similar elements

- Figures and tables
- Diagrams
- Graphs
- Lists
- Algorithms
- Examples and definitions
- Equations

Figures and tables

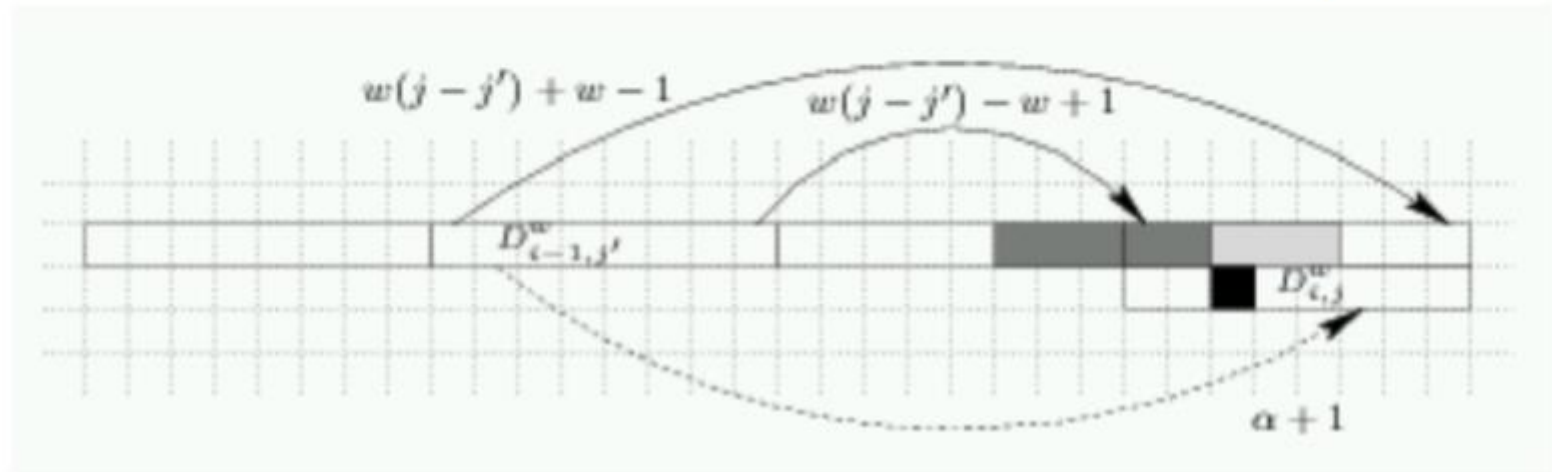
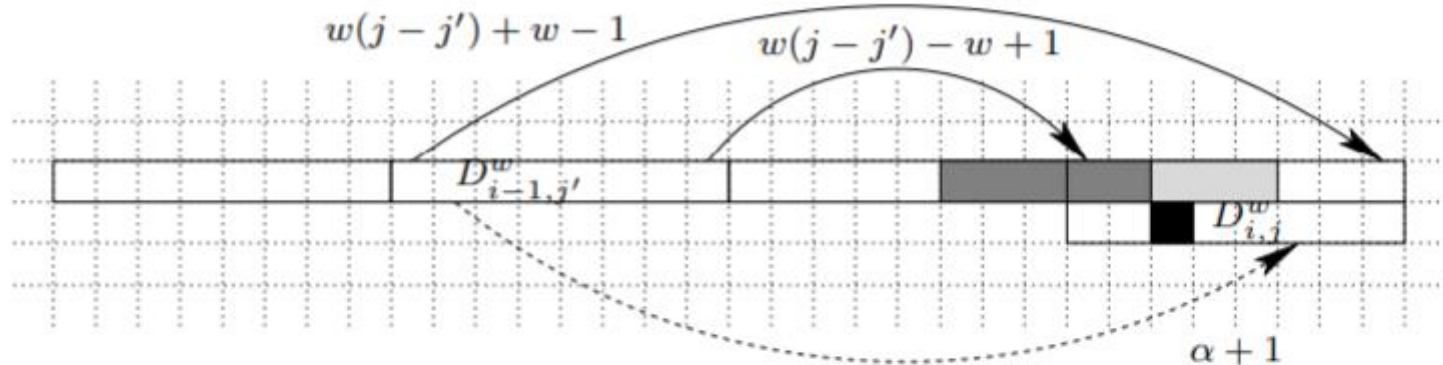
- General rules
- Vector graphics
- Captions
- Expressions

General rules

- Figures illustrate the models or the results, and tables give summaries.
- Two rules for displaying figures and tables
 1. All figures and tables must be referred in the text.
 2. There is no sense to express trivial things as a figure or a table
- If there is no need to refer to a figure/table in the text, the figure/table is probably not needed!
- Avoid repeating the same data in several places.
- Discuss only the most important items of the table in the text.
- A figure should be easy to understand..
- If two tables/figures should be compared, position them next to each other.

Vector graphics

- Use vector graphics, not raster graphic (bitmaps)! There is a big difference in quality:



Tables

- Is a table the best way to present your information?
- Consider graphs, bar charts or pie charts.
- Dependent tables (small) can be placed within the text, even as part of a sentence.
- Independent tables (larger) are separated from the text with table numbers and captions. Position them as close as possible to the text reference.
- Complicated tables should go in an appendix.

Captions

- Each table or figure should be understandable by its own. Give a brief but clear explanation or a title in the caption.
- Explain all special abbreviations, symbols, special use of underlinings, dashes, parantehtses, etc.
- Use the same style in all tables.
- If you copy (draw again) a table or a figure from some other source, then give a reference to the original source in the end of caption.

Expressions

- When you refer to figures and tables you can use the following expressions:
- The results are summarized/reported in Table 1
- The results are represented in
- Figure 2 illustrates
- In the Figure we observe
- The model is given in Figure 7
- etc.

Diagrams

- Keep them simple.
- Draw them specifically for the report.
- Put small diagrams after the text reference and as close as possible to it.
- Think about where to place large diagrams.

Graphs

- Graphs must be used particularly when trends are shown or series of data are compared.
- Graphs containing data that need to be compared must be combined in the same figure so that the reader can make a direct comparison between the values.
- The quantity of information in one graph must be limited so that the different symbols can be clearly distinguished.
- The axes of a graph must be named in words, in conjunction with units.

Lists

- Lists are not separate objects, and they are introduced in the text.
- Use list only when they are necessary!
E.g. "The main criteria of X are (the following):"
 - Criterion 1
 - Criterion 2 – ...Or "The method consists of five steps:" + a list
- If you list only a couple of items, you can usually write them without a list. Use lists when the clarify things!

Algorithms

- Give only the main algorithms in the text, and in an appropriate abstraction level (pseudocode)
- Fix the pseudocode notation and use it systematically
- Simple methods can be described by a numerated list of steps
- Logical and set operations are often useful when you describe algorithms in an abstract level
 - E.g. for all $x_i \in X$, $T = T \cup \{p_i\}$, find such $S \subseteq T$ that $q(S), \dots$
- If you write longer algorithms, insert them into a figure or an environment of their own.”

Example of Algorithm

Alg. 3 $\text{dfs}(X, d, \min_f, \text{last})$. A depth-first search of the self-referent sets in subgraph $G' = (V', E')$.

Input: $X \subseteq V, d, \min_f, \text{last}$

Output: $Y \subseteq V'$

```
1      begin
2          if  $f_{ref}(X) \geq \min_f$  then
3              output  $X$ 
4          else if  $(f_{ref}(X) < 1 - \frac{d(1-\min_f)}{(|X|-1)\min_f})$ 
5              then return // search failed
6          for all vertices  $u \in V'$  ( $u > \text{last}$  and  $\exists v \in X (v, u) \in E$ ) do
7               $\text{dfs}(X \cup \{u\}, d, \min_f, u)$ 
8      end
```

Definitions

- A good definition
 - explains the defined concept.
 - is not a circular argument (where x is defined by y and y by x).
 - is not expressed by negative terms, if possible.
 - doesn't contain unclear, vague, or descriptive language •
 - defines only what is needed

Expressions for referring to a definition

- The definition of ... is the following:
- The definition of ... is as follows:
- Formally, we define ...

Equations

- Without equation numbers
- Remember the full stop in the end of the equation, if the sentence finishes!

“The prior probability of X is updated by Bayes rule, given new evidence Y :

- $$P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)}.$$
”

- If the sentence continues, then you need comma:
“The dependency is described by equation
< equation >,
where a is sg. and b is sg.”

Equations

- With equation numbers

$$P(X|Y) \frac{P(X)P(Y|X)}{P(Y)} \quad (3.1)$$

- If you want to refer to some previous equation, you have to give it a label like for examples.

- Only use mathematics where it is the most efficient way to convey the information.
- Longer mathematical arguments, if they are really necessary, should go into an appendix.

Mathematics (continued)

- Mathematical notation can be used to describe algorithms, data structures, automata, or any of the objects in computer science
- A lot of difficulties for readers if the mathematics is badly presented
- There are conventions of presentation for mathematics and mathematical concepts.

Theoremes

- Details of the proof can often be omitted in a paper.
- Theorems, definitions, lemmas, propositions and key examples should be numbered.
- If the proof need lemmas, put them at the right place for readability.
- When stating your proof in a paper, Use means to convey your argument with the greatest possible clarity

Readability

- Mathematics is usually presented in italics, to distinguish it from other text.
- If a displayed formula is sufficiently important it should be numbered,
- Notations should follow the conventions of your area rather than invent your own.
- Use appropriate brackets, braces, parenthesis
- Sentences should begin with a word, not digits or mathematics
- Do not use a long stream of complex expression
- Avoid unnecessary subscripts

Example

✗ Let $\langle S \rangle = \{ \sum_{i=1}^n \alpha_i x_i \mid \alpha_i \in F, 1 \leq i \leq n \}$. For $x = \sum_{i=1}^n \alpha_i x_i$ and $y = \sum_{i=1}^n \beta_i x_i$, so that $x, y \in \langle S \rangle$, we have $\alpha x + \beta y = \alpha(\sum_{i=1}^n \alpha_i x_i) + \beta(\sum_{i=1}^n \beta_i x_i) = \sum_{i=1}^n (\alpha \alpha_i + \beta \beta_i) x_i \in \langle S \rangle$.

Should be replaced with

✓ Let $\langle S \rangle$ be a vector space defined by

$$\langle S \rangle = \left\{ \sum_{i=1}^n \alpha_i x_i \mid \alpha_i \in F \right\}.$$

We now show that $\langle S \rangle$ is closed under addition. Consider any two vectors $x, y \in \langle S \rangle$. Then $x = \sum_{i=1}^n \alpha_i x_i$ and $y = \sum_{i=1}^n \beta_i x_i$. For any constants $\alpha, \beta \in F$, we have

$$\begin{aligned} \alpha x + \beta y &= \alpha \left(\sum_{i=1}^n \alpha_i x_i \right) + \beta \left(\sum_{i=1}^n \beta_i x_i \right) \\ &= \sum_{i=1}^n (\alpha \alpha_i + \beta \beta_i) x_i, \end{aligned}$$

so that $\alpha x + \beta y \in \langle S \rangle$.



Lecture materials derived from:

- Justin Zobel (2014), *Writing for Computer Science*, Springer.
- Wilhelmiina Häämäläinen(2006), *Scientific Writing for Computer Science Students*, Course material, University of Joensuu.
- Nicole Kelley (2006), Basics of Technical Writing, slides of MIT course.