Writing for

Computer Science & Engineering

Ki-Il Kim

Department of Computer Science and Engineering

CNU

Graphs, Figures, and Tables

Introduction

- Well-chosen illustrations breaths life into a paper, giving the reader interesting visual elements to browse and highlighting the central results and ideas.
- A typical figure consists of visual matter such as a graph or diagram, or of textual matter such as a table, algorithm, or, less commonly, complex mathematics
- Some information is best presented in a pictorial form, such as a graph or figure, to show trends and relationship. Other information is best as a table, to show regularities

Graphs

- Graphs are often the best way to present numerical results
  - Use graphs wherever appropriate to elegantly summarize numbers and make obvious the behavior and tends that you wish to demonstrate
  - If you must list the numbers as well, put a detailed table of results in an appendix, but in many cases the trend is the interesting outcome
- You should present information because it is supporting evidence for a hypothesis, not because it is an output of some program
  - Don’ t flood you paper with statistics and avoid repetition
- Graph should be simple, with no more than a few plotted lines and a minimum of clutter
  - If the graph is being used to demonstrate variation in output values for a range of input values, the horizontal or x-axis should be used for the parameter being varied
  - Plotted lines of discrete data should always have points marked by distinctive marks such as circle, boxed, or triangles
  - Consider using greys, colors, or line thickness rather than dots and dashed to distinguish between lines

Graphs

- Colors are more eye-catching than are greys, but, in a graph, do not necessarily communicate better
  - Some journals do not print in color
  - Colors can render inconsistently in different media
  - Greys are emotionally neutral and don’ t carry the subconscious messages that colors can
- Minimize use of unnecessary elements and remove all decoration
  - Is a legend necessary?
  - Do the captions have to be in a large font
  - Are the fonts and font sizes different to those of the rest of the paper
- Great care is needed to ensure that graphs are effective at communicating results

Graphs

- Example
  - The graphs are rectangular rather than square with the legends placed in spare space within the body of the graph
  - Legend needs to be placed where it can’ t be confused with other material

Graphs

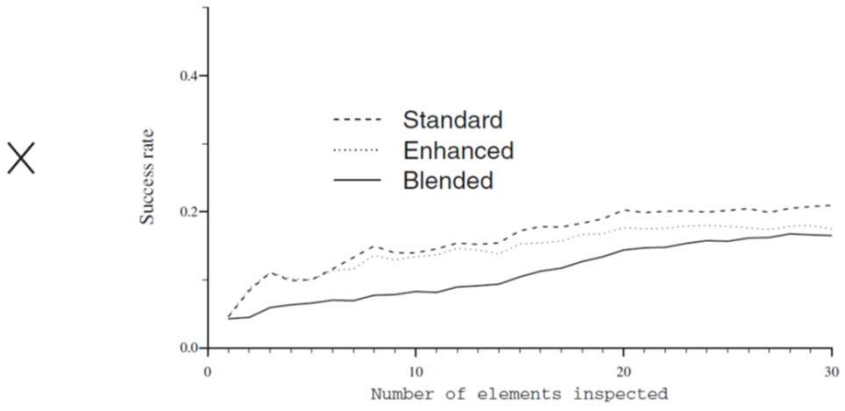


FIGURE 7. *Success rate as the number of inspected items is increased. It is clear that blending is not effective.*

- Poor use of vertical space
- Awkwardly placed legend
- Changed font and size
- Inconsistent with the main text

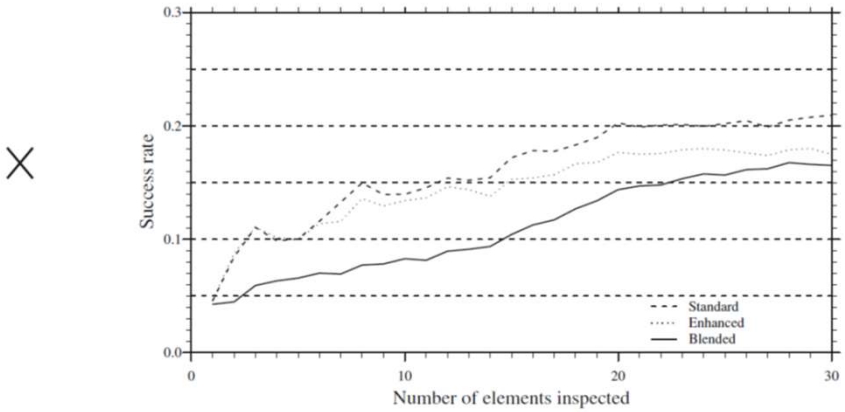


FIGURE 7. *Success rate as the number of inspected items is increased. It is clear that blending is not effective.*

- Partially corrected vertical scaling and fonts
- Unnecessary ornamentation
- Too small font size
- Grid line and heavy border

Graphs

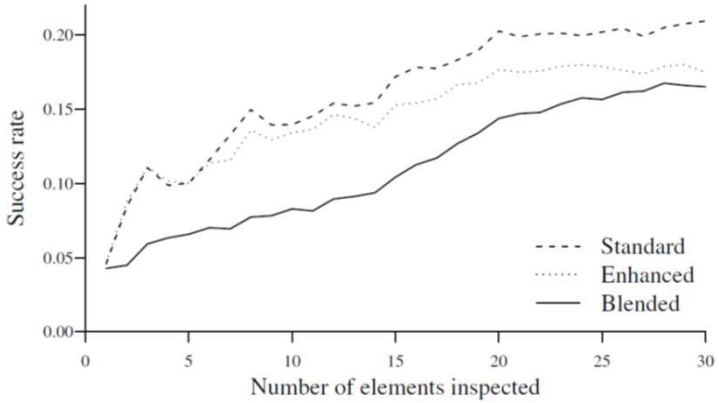


FIGURE 7. *Success rate as the number of inspected items is increased. It is clear that blending is not effective.*

- Corrected vertical scale
- Removed unnecessary tick mark

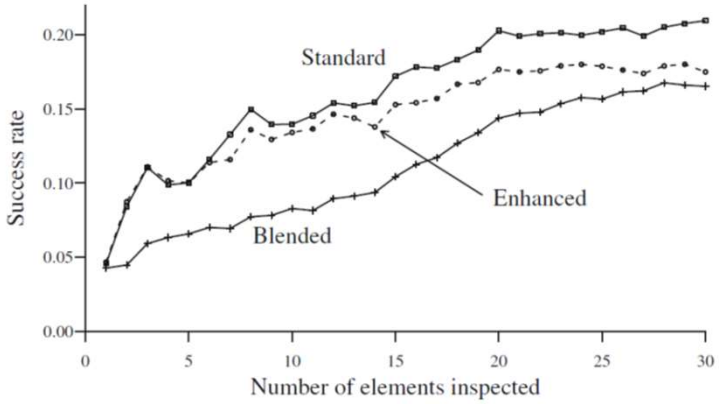


FIGURE 7. *Success rate as the number of inspected items is increased. It is clear that blending is not effective.*

- Stronger data lines
- Directed labelling
- Line ticks

Graphs

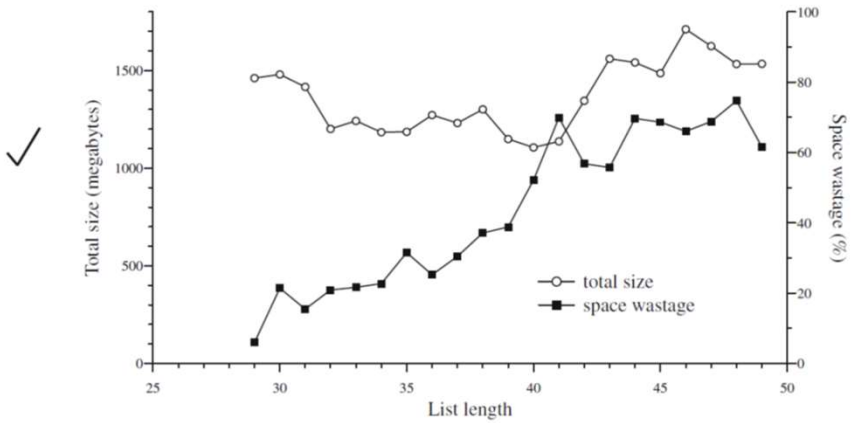


FIGURE 2. *Size and space wastage as a function of average list length.*

- Two functions
- Label the axes

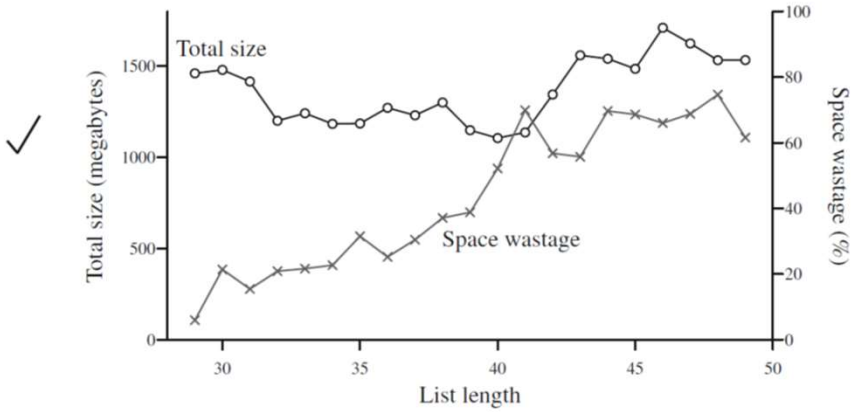


FIGURE 2. *Size and space wastage as a function of average list length.*

- Removing distracting elements
- Label the axes

Graphs

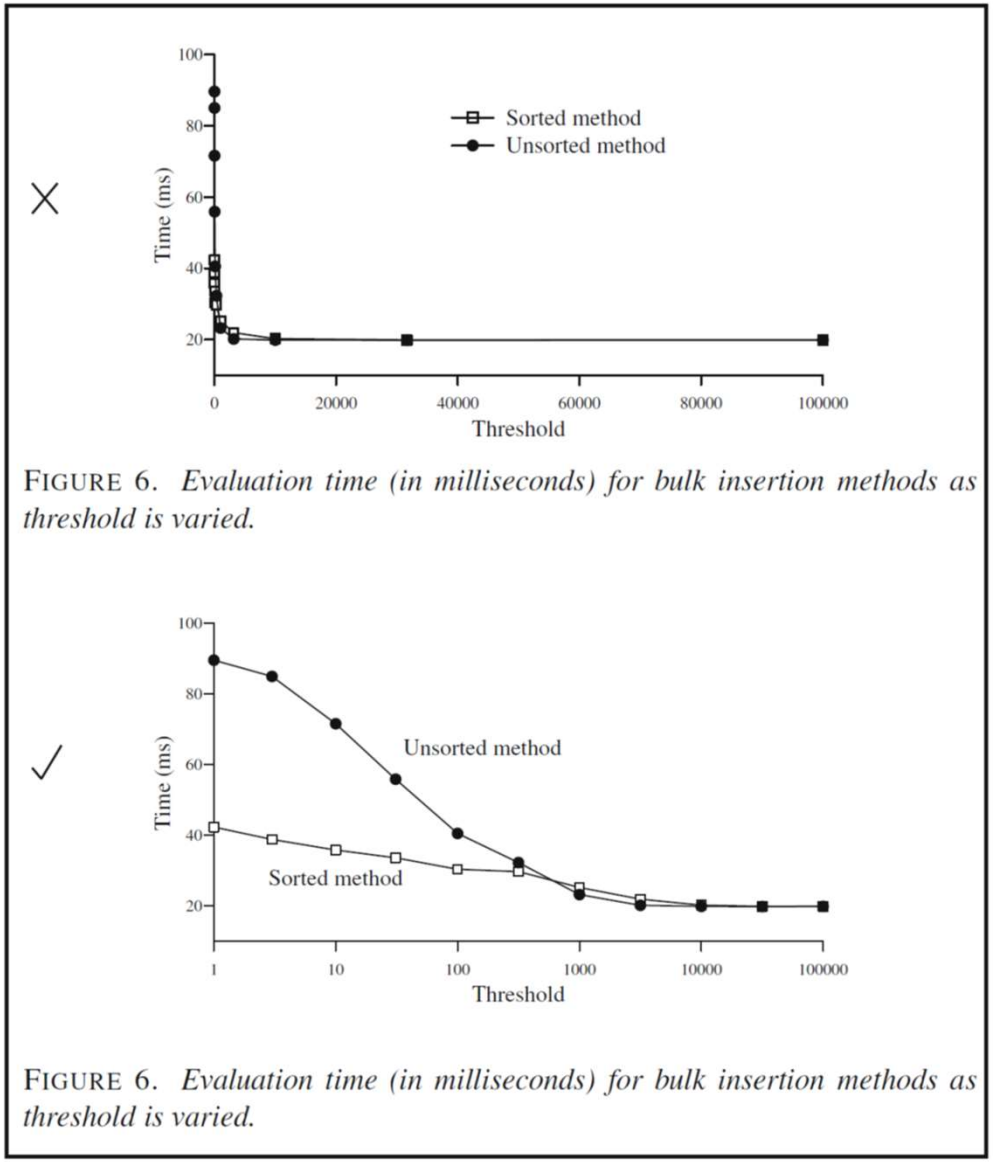


FIGURE 6. *Evaluation time (in milliseconds) for bulk insertion methods as threshold is varied.*

FIGURE 6. *Evaluation time (in milliseconds) for bulk insertion methods as threshold is varied.*

- Logarithmic scaling

Graphs

- Logarithmic axis
  - Useful because they show behaviors at different orders of magnitude
  - Useful when plotting problem size against algorithm running time, as different asymptotic growth rates give straight lines of different slope
- Log scaling is not always appropriate
  - If one algorithm is 30% faster than another at all scales, then, depending on overall scale, their performance could be almost indistinguishable on a log–log graph
- In some cases, data that seems innately tabular can be represented as a graph
  - Often a bar graph is suitable because the items being compared are not ordered
  - Such data should not be represented by a line chart in which the points are connected into a continuous line, which would imply that the axes were related by a function

Graphs

Data set	Method	
	A	B
Small, random	11.5	11.6
Large, random	27.9	17.1
Small, clustered	9.7	8.2
Large, clustered	24.0	13.5
All documents	49.4	60.1
First 1000	21.1	35.4
Last 1000	1.0	5.5

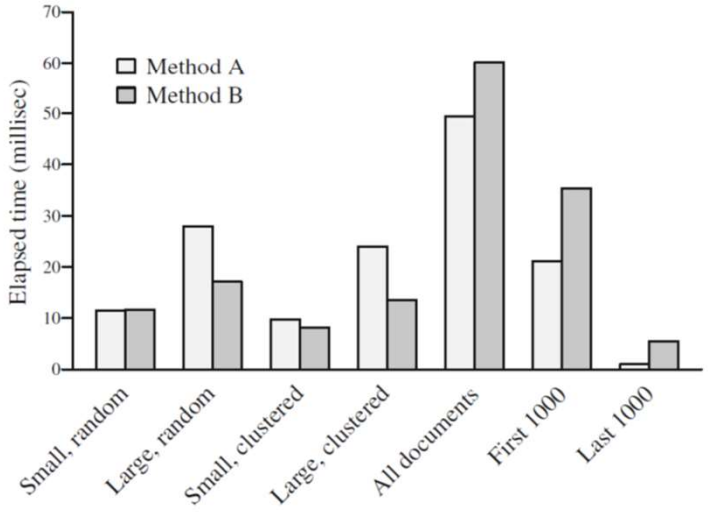


FIGURE 2. *Elapsed time (milliseconds) for methods A and B applied to data sets 1–7.*

- Graph is a better choice

Graphs

TABLE 8.4. *Tradeoff of space against time for methods A to G.*

Method	Space (%)	Time (ms)
A	1.0	7,564.5
B	31.7	895.6
C	44.7	458.4
D	97.8	71.8
E	158.1	18.9
F	173.7	1.4
G	300.0	0.9

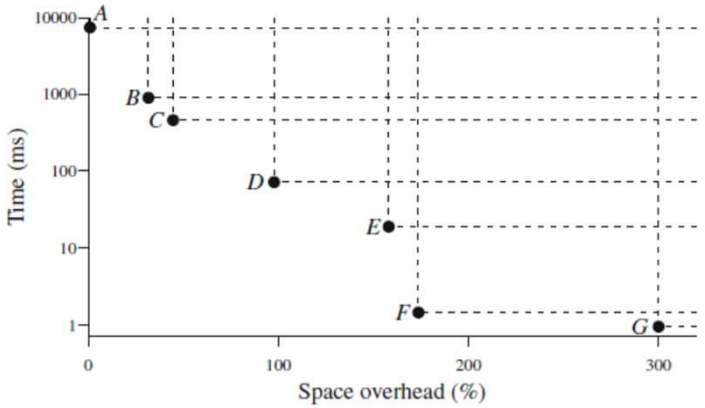


FIGURE 8.4. *Tradeoff of space against time for methods A to G. The boxed area to the right and above each point is of unacceptable performance: any method in that area will be less efficient with respect to both space and time than the point at the box's corner.*

- Difficult to interpret

- Illustrates the frontier of efficiency
- Describe the region that would represent an improvement on existing methods

Graphs

- Graph-drawing tools allow bar graphs to be three dimensional, but the addition of depth is deceptive; if one bar is twice the height of another, the depth exaggerates the difference
- Graphs are used to illustrate change in one parameter as another is varied
  - If two parameters, say B and C, depend on another, A, then a good solution is to plot A on the x-axis and have two y axis, one for each of B and C
  - If two parameters, say D and E, jointly determine a third, F, in some complex way-thus describing a three-dimensional space-the problem is more difficult. Use of a three-dimensional graph is an option -> such graphs can be powerul explanatory tools, but should not be used merely because they are dramatic or eye-catching
  - Plot D against F several fixed value of E, and use these results to choose an E value that yields a representative graph

Graphs

- Where several methods of achieving the same aim being illustrated, the axes in each graph should have the same scale
  - If you are comparing different data structure and a separate graph is used for each one, the axes should be consistent from one graph to the next -> direct comparison between the mothed
  - Comparison is easier with several lines on one graph
- For a quantity that is wholly positive and measured on a ratio scale, the Y axis should start from 0, as starting form a higher value can exaggerate of change of difference. If the axis is started at a higher value for clarity, the reader should be alerted to this in the caption

Graphs

- Valuable features of software packages for drawing graphs
  - Placing of several lines on one graph
  - A range of symbols for marking points
  - The ability to crate custom symbols of custom size for marking points
  - …….
  - Page 166

- Good software packages for drawing graphs
  - Gnuplot
  - Origin
  - …….

Diagram

- Diagram serve many purposes in computing papers
  - Illustrate processes or architecture
  - Explain data structures and algorithms, present relationships, visualize data, and show examples of interfaces
  - ER-diagram, automata
- To design a diagram that is to be created with a manual tool
  - First step is often to do initial sketches by hand
  - Well-proportioned, makes good use of the space, laid out well
  - Doesn't have the elements bunched to one side
- Diagram
  - A diagram should not be too dark
  - Lines should not be too heavy
  - Pictorial elements should be used consistently
  - Color can be highly effectively
  - Lines should not touch each other

Diagram

- Diagram
  - A diagram should not be too dark; keep it as sparse as possible
  - Use meaningful labels and make the point size and fond of the labels similar to that of other text
  - There should be no more than two or three fonts and font sizes
  - Lines should not be too heavy
  - Pictorial elements should be used consistently (arrows and lines of the same kind have the same meaning)
  - Color can be highly effectively
  - Lines should not touch each other

- Diagram, like graphs, can add greatly to the clarity of paper. But, be aware that the design of good diagram is not easy.
- Expected to revise your picture as often as you would your writing

Diagram

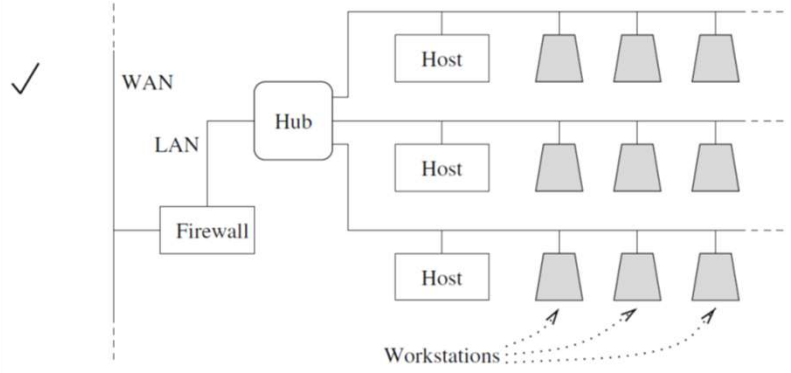


FIGURE C. *Revised network, incorporating firewall and hub with hosts and workstations on separate cables.*

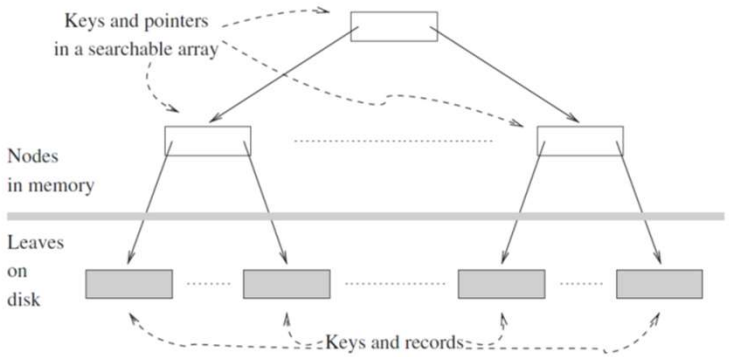


FIGURE 1.3. *Tree data structure, showing internal nodes in memory and external leaves on disk; omitted nodes are indicated by dotted lines. Nodes allow fast search and contain only keys and pointers. Leaves use compact storage and contain the records.*

- Shading and dashing
- Entities are of the same kind

Diagram

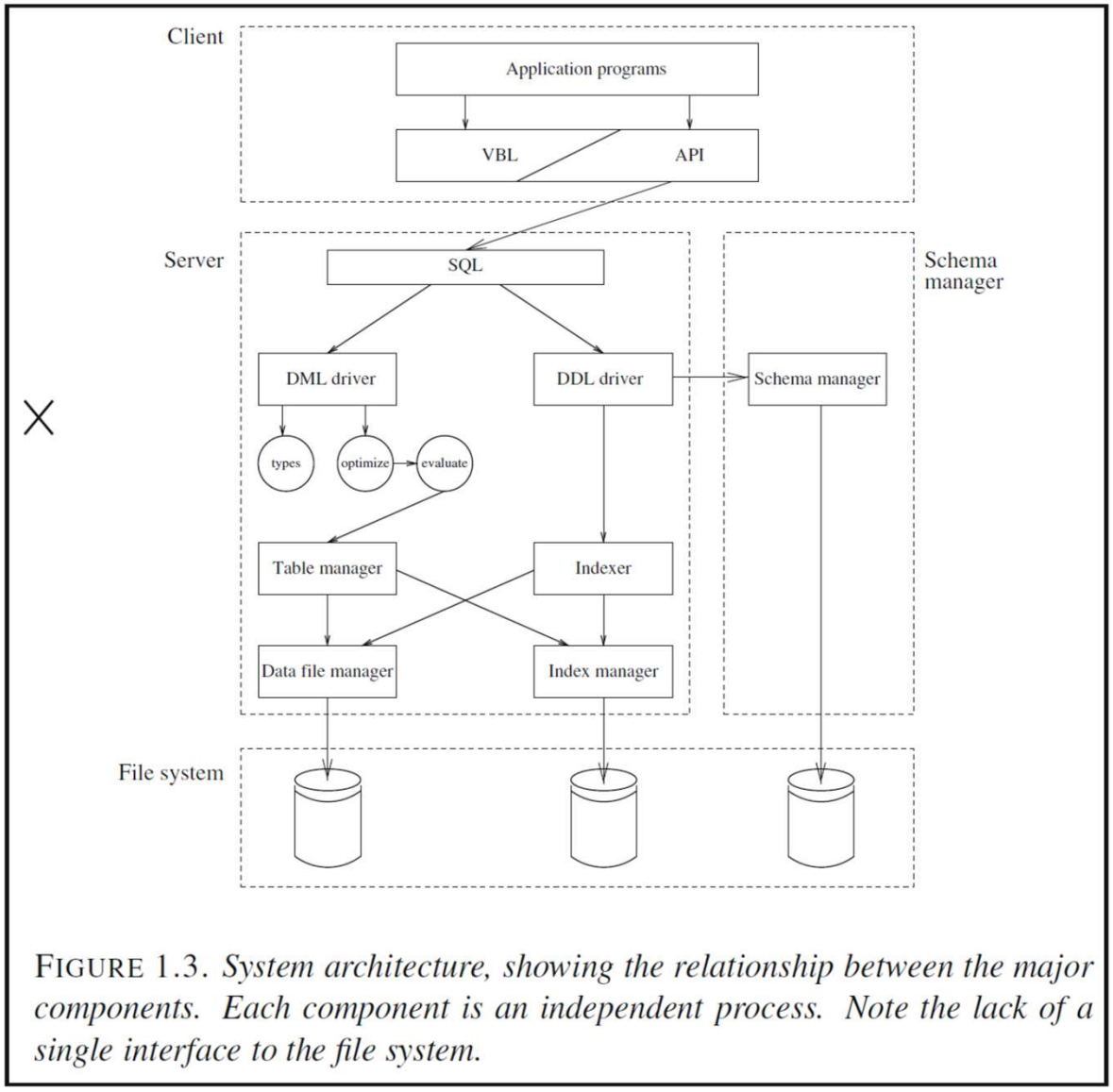


FIGURE 1.3. *System architecture, showing the relationship between the major components. Each component is an independent process. Note the lack of a single interface to the file system.*

- Too much clutter
- Too small font
- Other details are more emphasized
- Some of the internal detail should be omitted
- Arrow add little information
- Arrow should point both ways

Diagram

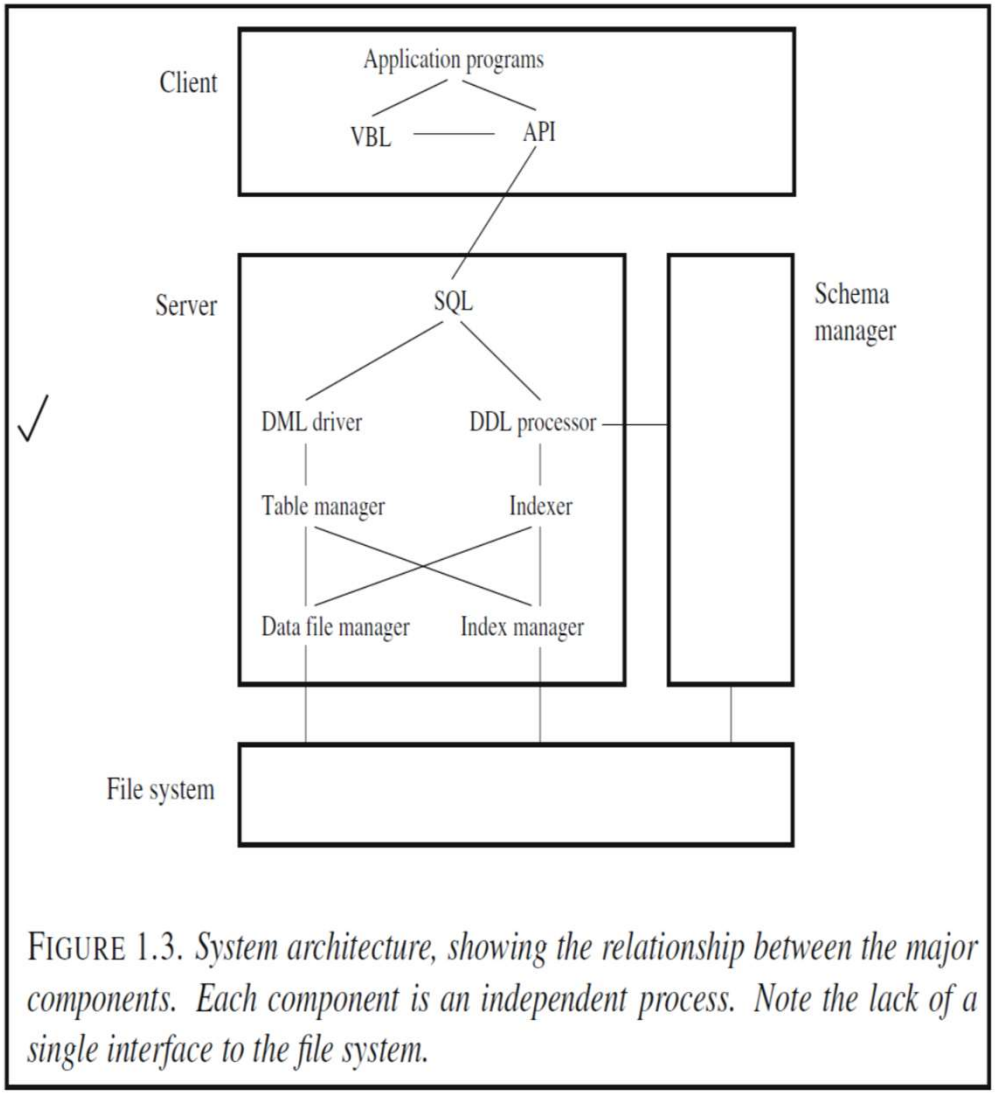


FIGURE 1.3. *System architecture, showing the relationship between the major components. Each component is an independent process. Note the lack of a single interface to the file system.*

- Overall structure is more prominent
- Discarded some minor features
- Removed unnecessary inner boxes
- Use of sharing : further improvement

Diagram

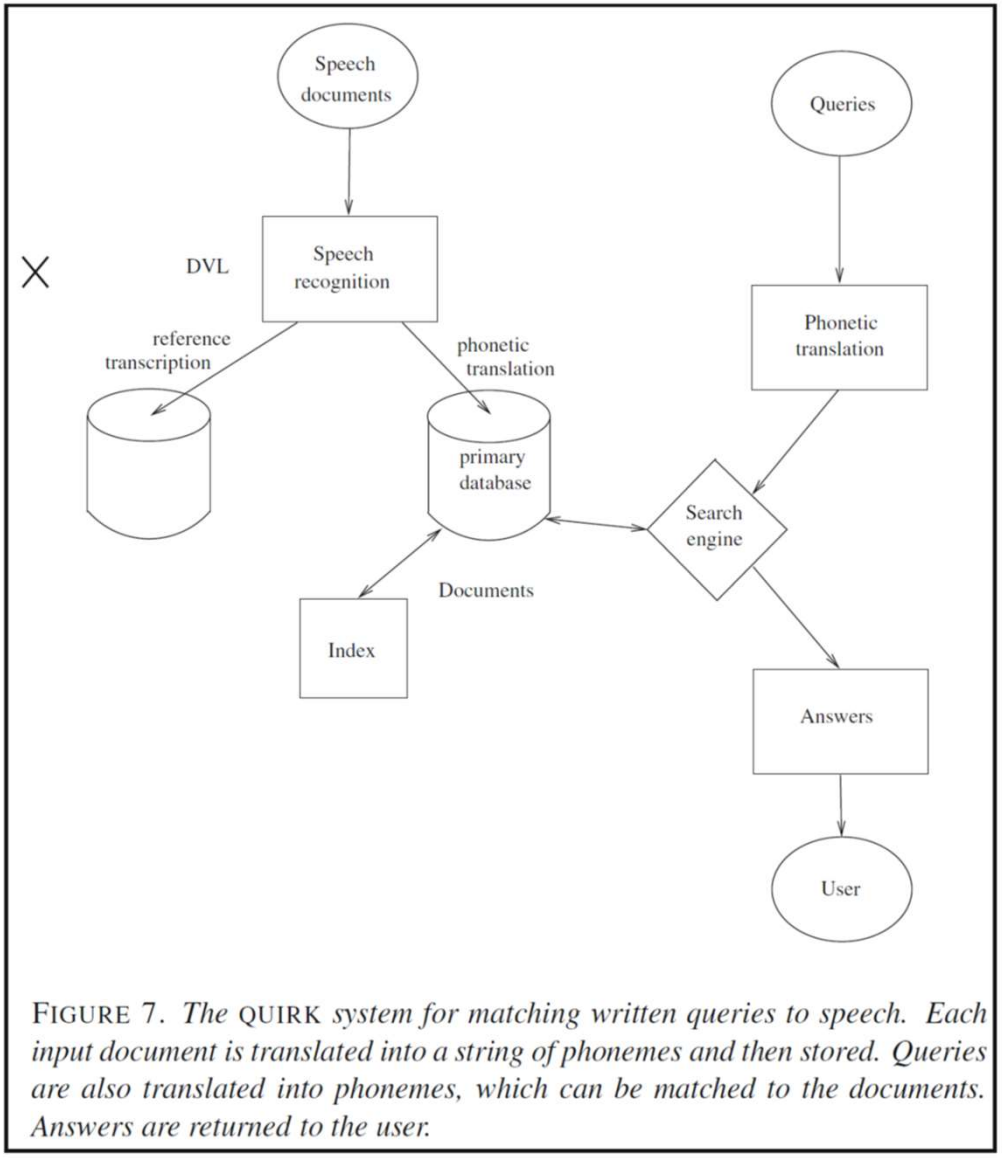


FIGURE 7. *The* QUIRK *system for matching written queries to speech. Each input document is translated into a string of phonemes and then stored. Queries are also translated into phonemes, which can be matched to the documents. Answers are returned to the user.*

- Disorganization
- Inconsistent element
- Data is in both ovals and boxes
- Arrowheads touch other lines
- Creating messy intersection

Diagram

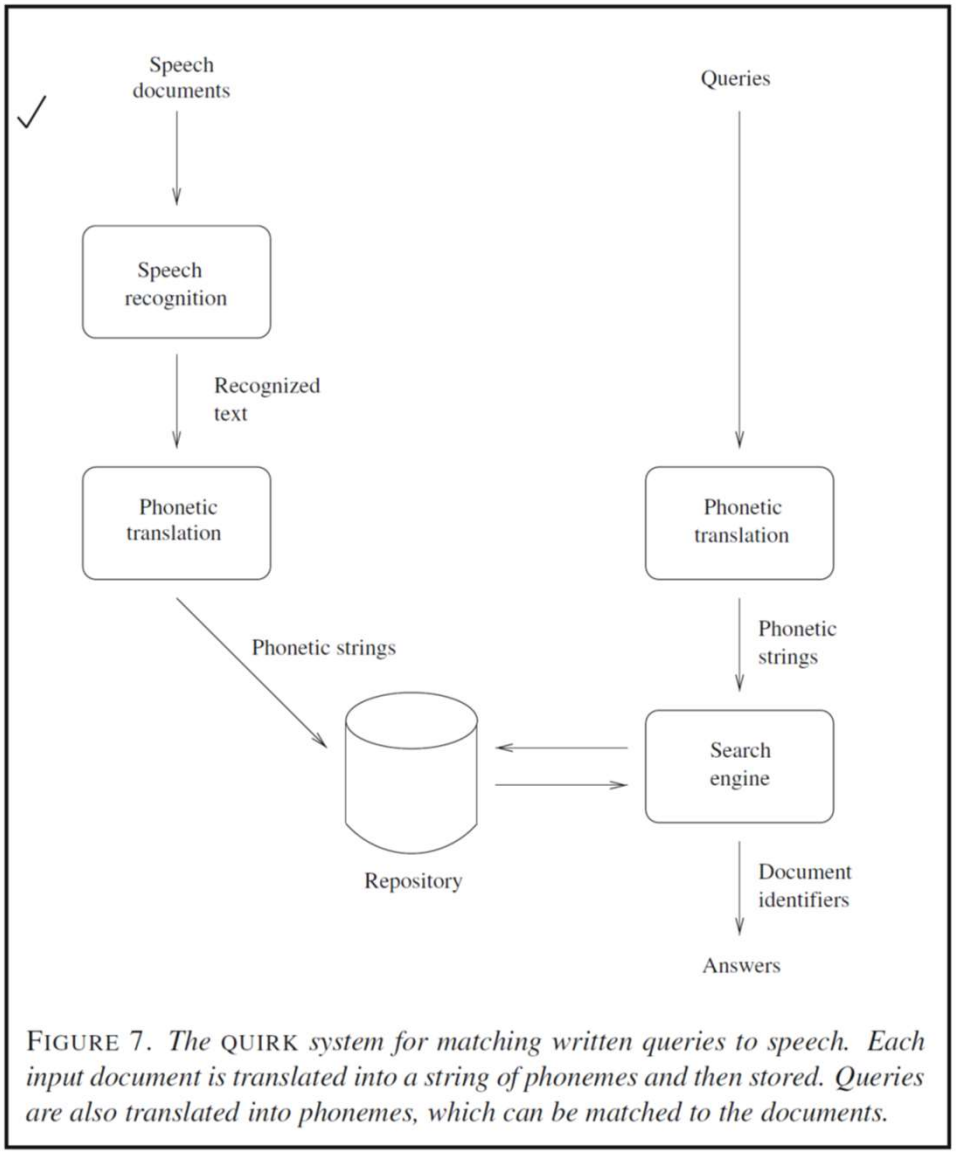


FIGURE 7. *The* QUIRK *system for matching written queries to speech. Each input document is translated into a string of phonemes and then stored. Queries are also translated into phonemes, which can be matched to the documents.*

- Parallels between document processing
  and query processing are emphasized
- Removed unnecessary material
- Two-headed arrow is replaced by two arrows, to shows that data is exchanged

Diagram

Copyright

- Figures from another source can only be reused with permission of the author and the publisher of the original
- You may also need to include the original copyright statement

More example

- Refere to page 168 ~ 171

Tables

- Tables are used for presentation of information that is unsuitable for graphs and figures, such as the properties of each of a series of datasets or data where the exact values are important
- Well-designed table
  - Logical hierarchical structure by separating rows and columns in double lines

TABLE 6. *Statistics of text collections used in experiments.*

STATISTICS	SMALL	LARGE
Characters	18,621	1,231,109
Words	2,060	173,145
After stopping	1,200	98,234
Index size	1.31 Kb	109.0 Kb

TABLE 6. *Statistics of text collections used in experiments.*

	Collection	
	Small	Large
File size (Kb)	18.2	1,202.3
Index size (Kb)	1.3	109.0
Number of words	2,060	173,145
— after stopping	1,200	98,234

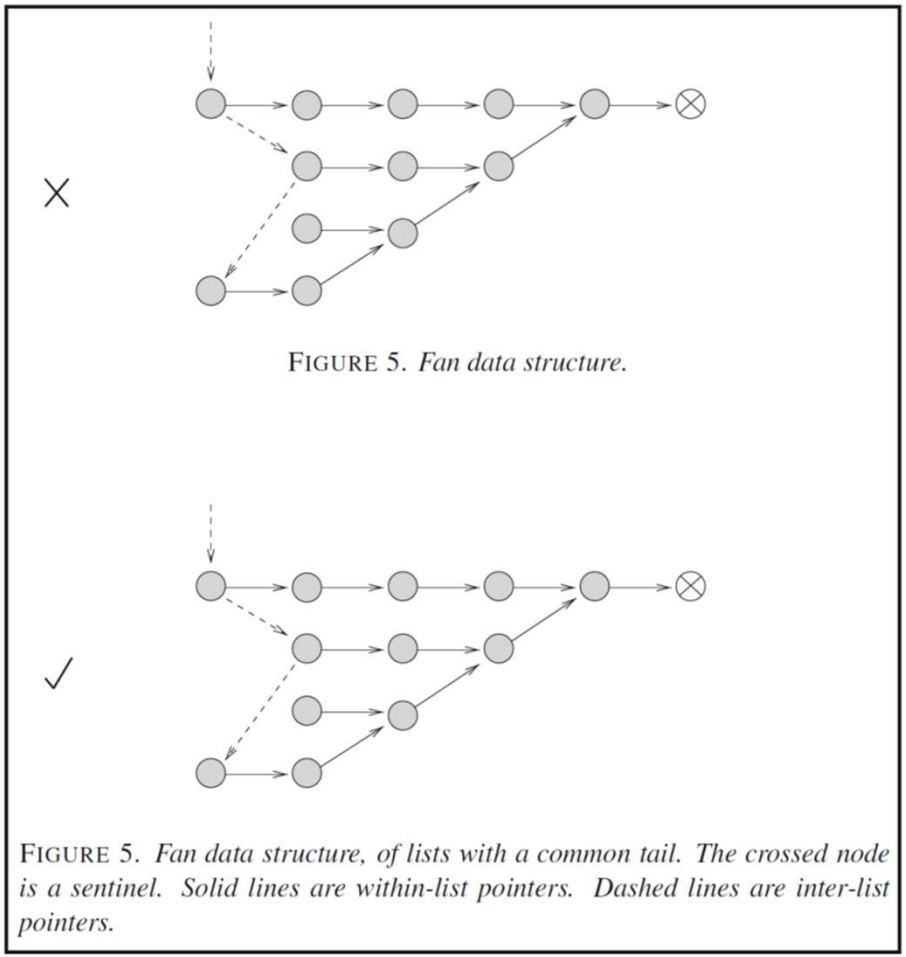
Tables

- Table
  - Item below a column head should be of the same kind or about the same thing
  - Should be open and uncluttered with ample with space
  - Don’ t have too many horizontal or vertical rules
  - Don’ t make tables too dense
  - No position should be blank; if there is no applicable value, put in a dash, and explain somewhere what it means
  - Numbers should aligned on the decimal point

- Understanding a table of any complexity is hard work. For presentation of results, graphs or explanatory text are preferable; have a table to which the interested reader can refer, but don’ t rely on a table to convey essential information

Captions and Labels

- Captions and labels should be informative
  - Fully describe the figure’ s major element
  - The caption is usually placed below a figure, but above a table



- Less dependent on the paper’s text

Captions and Labels

- Each figure and table should be numbered to allow easy reference
  - A figure is usually at the top or the bottom of a page, or on a page by itself, to set it apart from ordinary text
  - A figure or table should always be introduced and discussed in the main text, preferably just before or on the papge on which it oocurs
  - If you don’ t have anything to say about a figure or table, leave it out

Axes, Labels, and Headings

- Space constraints on axes, labels, and headings may mean that some terms have to be abbreviated

  ✗ The abbreviations "comp.", "doc.", and "map." stand for "compression", "document", and "mapping table" respectively.

  ✓ The effect of compression on the documents and the mapping table is illustrated in the second and third rows.

- Where appropriate, units should be stated in labels, Size (bytes) rather than Size
- Some reader get confused by scaling on axes and labels
  - An axis is labelled as "CPU time (seconds x 10^{-2})
  - Convention : multiple axis values by 10^{-2}
  - Some readers may assume that the axis values have already been multiplied by 10^{-2}
  - It is helpful to include some representative numbers in the text, because graphs are hard to read with any precisions
  - Figure 5 shows how time and space tradeoff as node size is varied: as con be seen, response of under a second is only possible when size exceeds 11Kb.

Writing for

Computer Science & Engineering

Ki-Il Kim

Department of Computer Science and Engineering

CNU

Editing

Introduction

- The writing of a paper begins with a rough draft, perhaps based on records of experiments or sketches of a couple of theorems
- Next phase usually consist of filling out the draft to form a contiguous whole: explaining concepts, adding background material, arranging the structure to give a logical flow of idea
- Finally, the paper is polished by correcting mistakes, improving written expression, and taking care of the layout
- Styling of the paper is the last phase and has the most impact on a reader. It should not be neglected

Introduction

- Few writers are good at judging their own writing
  - Discovery of shortcomings in your text takes time and effort
  - Careful reading, a willingness to admit to mistakes, the regrets of discarding text that was hard to create
  - The difference between a weak writer and a strong writer is not ability to write fluently, but the effort taken to diligently edit and revise

Consistency

- Editing is the process of making a document ready for publication or examination
  - Much of editing consists of checking the document for errors that fall under the heading of consistency
  - Use the checklist when revising your paeprs, or when proofreading papers for others
- Some of checklist
  - Are all of the components present: title, authors, abstract, and so on?
  - Are the acknowledgements complete and accurate?
  - Is the ordering of material correct?
  - Are the titles and headings consistent with the content?
  - Have all terms been defined?
  - Is the style of definition consistent? For example, were all new terms introduced in italics, or only some?
  - Has terminology been used consistently?
  - .....

Consistency

- A surprisingly effective editing exercise
  - Pretend to be a reader
- Early drafts
  - Tend to be repetitive and long-winded
  - Not only are concepts awkwardly expressed and sentences unwieldy, but material on one theme might be in separate parts of the paper
  - It is common to find similar material included several times, in different palces, particularly when there are several authors
- Ordering may need to be reconsidered one the paper is complete
  - When material is moved from one place to another, check that the text in each location is intelligible and appropriate in the new context
- For many papers, editing leads to removal of text. Don't be afraid to shorten your papers: cutting will improve the quality. Edit for brevity and balance. Omit or condense any material whose content or relevance to the paper's main themes does not justify its length

Style

- Another kind of editing is for style and clarity, and is perhaps the hardest part of finishing a paper
  - When revising the text of other writers, it is often preferable to make minimal changes: correct the presentation but retain the flavor of the original text
  - Don’ t expect to impose your style on someone else
- Most journals have a preferred style for elements such as references, figure numbering, spelling, table layout, and capitalization. If you are planning to submit to a particular journal, consider using its style.

Proofreading

- There is no excuse for a report that contains spelling errors
  - Displaying not only your inability to spell, but also your casual attitude to your work
  - Find a spell checker
- A common error is to instead type some other word that shares a few initial letters
  - Being –> begin, form –> from, relation –> relative, compute –> complete, two –> too
- Identify and look for your own common errors
  - Incomplete sentences and sentences that have been run together inappropriately
  - Check for errors in tense and in mumbers in the use of plural and singular forms
  - When you identify an error that your often make, add it to a checklist, and look for it whenever you revise
  - Read each sentence carefully and ask yourself how easy it is to understand

Proofreading

- It is particularly important to check the bibliography.
  - Reader will use it to track down reference, so any garbling of information can lead them astray, and other writers may be offended if you have misreferenced their papers
  - Format should be consistent and each reference should include enough information to allow readers to locate it
- Always get someone else to read your work before you submit it or distribute it
  - Misunderstood a relevant articles, or made a logical error
  - Most authors are poor at detecting ambiguity in their own text
  - If something has been misunderstood, the paper needs to be changed, although not necessarily in the way proofreader recommends

Choice of Word-Processor

- The choice is dictated by availability, but also how well the available word-processors cope with the demands of authoring
- Two kinds of word-processor
  - Visual style – Word
  - Compiler style – Latex
- There are many circumstances in which I choose to use a visual word-processor, but technical writing is not among them.


Q
&
A