Scholarship Skills

Andrew P. Black

Version of 2015-01-07

Contents

1	**Reading**	**1**
1.1	Introduction	1
1.2	Why Read?	1
1.3	Deciding What to Read	2
1.4	The Mechanics of Reading	3
1.4.1	Establishing context	3
1.4.2	Goals of the first pass	4
1.4.3	Goals of the second pass	4
1.4.4	Goals of the third pass	5
1.4.5	Goals of the fourth pass	6
1.4.6	Remembering what you've read	6
1.5	Cracking a tough paper	6
1.6	Organizing Source Materials	7
1.7	Chapter Summary	8
2	**Background Research**	**9**
2.1	Introduction	9
2.2	Why Do Background Research?	10
2.3	Why is Background Research Challenging?	11
2.4	Where to Search	11
2.4.1	The Web	11
2.4.2	Ask Around	12
2.4.3	Important skills for doing a good literature review	13

2.5 Conducting a Literature Search . . . 13
2.5.1 Developing background knowledge . . . 14
2.5.2 Searching . . . 16
2.6 Supporting Technologies . . . 17
2.7 Chapter Summary . . . 17

3 Core Rules for Better Writing 19
3.1 Introduction . . . 19
3.2 What is Good Writing? . . . 19
3.3 The Core Rules of Writing . . . 20
3.3.1 Use Active Verbs . . . 20
3.3.2 Put Key Ideas in Lead Position . . . 22
3.3.3 Don't make Unsubstantiated Statements . . . 23
3.3.4 Be Concise . . . 24
3.3.5 Be simple . . . 24
3.3.6 Use a Consistent Lexical Set . . . 25
3.3.7 Define terms when first used . . . 25
3.3.8 Avoid one-word sentences and single sentence paragraphs . . . 25
3.3.9 Re-write . . . 25
3.4 Chapter Summary . . . 26

4 Mechanics 27
4.1 Basic Grammar . . . 27
4.1.1 Singular and Plural . . . 27
4.1.2 Past and Present Verbs . . . 28
4.1.3 Articles . . . 28
4.2 Spelling . . . 29
4.2.1 Word Pair Confusables . . . 29
4.2.2 Counting and Mass Nouns . . . 30
4.2.3 Parallel Clauses . . . 32
4.3 Clear Writing . . . 32
4.3.1 Word Choice. . . . 32
4.3.2 Word Placement . . . 33
4.3.3 Non-referential Pronouns . . . 34

4.3.4 Lists . . . 34
4.3.5 Use of Jargon . . . 35
4.3.6 Use of Latin . . . 35
4.4 Style and Notation . . . 36
4.4.1 Citations . . . 36
4.4.2 Numbers . . . 36
4.4.3 Decimal points . . . 36
4.4.4 Capitalization . . . 37
4.4.5 Fonts . . . 37
4.5 Punctuation . . . 37
4.5.1 Quotations . . . 37
4.5.2 Colons and Semicolons . . . 37
4.5.3 Commas . . . 37
4.5.4 Solidus and Dashes . . . 37
4.5.5 Parenthesis . . . 38
4.6 Further Reading . . . 38

5 Citing your Sources 39
5.1 Introduction . . . 39
5.2 When to Cite References in Scientific Research Papers . . . 40
5.3 Citing References in your Text . . . 41
5.4 Formatting the Reference List . . . 43
5.4.1 Books . . . 43
5.4.2 For Articles from a Compilation. . . . 43
5.4.3 For an Article from a Journal or Magazine . . . 44
5.4.4 Internet sources . . . 44
5.5 Automating the Bibliography . . . 45

6 Presentations 47
6.1 Introduction . . . 47
6.2 Preliminaries . . . 48
6.3 Kinds of Talk . . . 49
6.4 Gather Material . . . 49
6.4.1 How much material do you need? . . . 50

6.4.2 Where do you find material? 50
6.4.3 Kinds of Material[2]. 50
6.5 Create the slides 51
6.5.1 Title and Introduction 51
6.5.2 Conclusion and Acknowledgments 52
6.5.3 Managing Time 52
6.5.4 Conveying the Key Ideas 52
6.5.5 Directing Attention 53
6.6 Practice the talk 54
6.7 Giving the Talk 55
6.7.1 Before the talk 55
6.7.2 Getting started 55
6.7.3 Grab your audience 56
6.7.4 Use the Available Space. 56
6.7.5 Directing attention 56
6.7.6 Using notes 57
6.7.7 Timing 57
6.7.8 Mistakes 58
6.8 Answering questions 58
Bibliography 60

[2]This part is based on the book "Business Communication" by Madhukany Jha

Chapter 1

Reading

If I have seen further,
it is only by standing on the shoulders of giants.

Isaac Newton, letter to Robert Hooke,
5 February 1675 os.

1.1 Introduction

As someone who plans to contribute to the sum of human knowledge, you are likely to spend more time reading than writing. Why? Because you have to be aware of the current state of your field before you can contribute to it, and because your own work must be firmly grounded on the work of others. Reading can be a powerful way of improving your own writing. Do you really like the organization, phraseology, typography or style of a paper that you are reading? Use this to improve your own writing. Is there something about an article that makes it hard to follow, difficult to comprehend, or just drives you crazy? Don't do that in your own work!

1.2 Why Read?

There are various reasons to read an article, and the way that you read will depend on your motivation. Are you looking to answer a particular question, trying to discover if your own work is novel, situating your work in a field, or simply educating yourself? You may also have been asked to review the article, write an abstract of it, or prepare a presentation on it; you may be looking for inspiration, or you may just be stuck on a bus and have nothing else at hand!

What might you get out of the reading process? My colleague James Hook shared the following list:

- The context of the paper.
- The thesis being investigated.
- The contribution.
- The method of investigation.
- The "power" of the results.
- The influence of the paper.
- The applicability of the results.
- Summary of the technical development.
- Details of any examples.

Your reason for reading the article will help you prioritize some of these benefits above others, and tell you where to focus your energies. If you don't know why you are reading an article, how can you know whether reading it will be worthwhile?

1.3 Deciding What to Read

When you select a topic, you may find yourself overwhelmed by materials from a wide variety of sources. How does one choose where to start? Some suggestions include:

- Ask your advisor or senior students about what conferences and journals are reputable in your area.
- Take a look at your advisor's web page to see where she is publishing.
- Pay attention to the reference sections of papers to see what conferences and journals are frequently mentioned and what papers are frequently cited.
- If possible, find surveys of your topic to read at first.

Once you have found a good paper, take a look at the references to expand your reading list. Sometimes, even a bad paper can have good references. In addition, the related work section is often a good source of background information on the context of the paper. Identify the research group that produced the paper and look at their publication list or annual reports to expand your reading list.

When you find a paper through a search engine, you will probably first approach it from its title. Read the title and ask yourself "what kind of paper

is this?" The title may give you some hint to infer the type or the purpose of the paper. Sometimes, the title will indicate the methodology used in the research, such as "Using bi-simulation to determine equivalence of DFAs". Some titles suggest that the paper is a comparison work, such as "Four types of . . . ". In general, titles provide a summary statement that either describes the research or presents the main conclusion.

Consider the credibility of the paper. What organization is the author affiliated with and what bias might that bring to the paper? What about the reputation of the conference or journal that published the paper? If the paper has been published for a while, has it been cited often? If the paper is too old, try to determine if it might have been superseded.

1.4 The Mechanics of Reading

Some people there are who, being grown,
forget the horrible task of learning to read.

John Steinbeck. The Acts of King Arthur and his Noble Knights. Pan Books, 1979.

Find a quiet place to read where you will not be disturbed or distracted. One of the reasons that I don't like reading at a computer is that it presents a constant source of distraction; find a situation that works for you. It is important to have some way of taking notes, electronic or on paper. Notes are important because they boost your understanding. If you highlight major points in your notes, you will not only have a handy reference if you need to summarize the article, but also you will have to decide what the major points *are*, and that will help you maintain focus. Notes will also help you in the future. When you have just read an article, you may understand it well, but it may not be so clear in a week or a year. So take notes, and file them with the paper.

Reading is more than just passing your eyes over the words in consecutive order. Reading is understanding. You can maximize understanding by tackling a book or article in several passes, sometimes spread over time. How many passes? Over what time period? These depend on how difficult the paper is, how sharp you are feeling on a given day, how familiar you are with the topic, and on how thoroughly you need to understand the material. I'll describe one useful approach as an example.

1.4.1 Establishing context

Even before you start reading a paper, it is important to build a little context. Take a look at the authors' name, the group they belong to, and the year in

which the paper was published. Knowing a little about the authors will help you understand their motivation a little better. A fairly recent paper gives you a technical summary of the research in that field. Conversely, if you find the background overwhelming, it may be better to pick an older paper cited in the references (perhaps from the same authors) to build up some expertise.

1.4.2 Goals of the first pass

On the first pass you will be skimming; skimming gives you an overview of the authors' work and helps you establish a conceptual map of the different parts of the paper. This map helps you make sense of the different sections of the paper and how they relate to each other. As a result, the process of understanding the paper will be much easier than starting off with a full read. Additionally, when you skim a paper you save time as you may decide to stop reading if the paper is not interesting to you. After skimming, whether you decide to continue reading or not, you now already know what the paper is about, which will increase your knowledge of the field.

On the first pass, it's probably sufficient to read the abstract, look at the pictures, and skim the list of references. The abstract will help you to determine if the paper is relevant to your goals, and to find out what kind of paper it is: a review, a minor incremental result, a landmark paper in a field, or a paper that started a new field. The "pictures" — the tables, graphs and diagrams — should give you a clue as to the major concepts and results. You should also see if you recognize the cited references or know what they are about — if not, you will probably have to do some background reading if you need to understand everything in the article in front of you.

1.4.3 Goals of the second pass

On the second pass, read the introduction, look at the start of each section or chapter, look at the examples, and read the summary or conclusion. The introduction should help you to understand the organization and the context of the article, and to decide if the authors are good writers. It took me a long time to figure this out, but I finally realized that there are more interesting papers out there than I will ever be able to read. Since I have to be selective, why not select the papers that are well written: the ones where the authors have made my task as a reader easy? The second pass should help you to decide how much work a more thorough reading will be, and whether the payoff is likely to outweigh the effort. On this basis, you might stop here, or you might decide to read only parts of the book or article.

Sometimes, authors will try to hide the fact that they are citing their previous work. For example, "Several researchers have recently identified the need to reduce data analysis time for very large data processing tasks [4, 8,

14, 15, 22, 23, 28]."[1] In this vaguely-worded sentence the co-authors wrote references 4 and 15. If the paper has been published for a while, checking how many times it has been cited can help to establish the importance of the paper, and help you determine if it is worth another pass.

1.4.4 Goals of the third pass

The third pass is a full reading, and you will probably find it useful to take notes as you read. You are rarely reading an article in isolation: you are more often building up your own knowledge of a particular research area. So try to capture the main contributions of the article, and how it is similar to and distinguished from other work.

During the full read, mark definitions of unfamiliar terms in the margin. Make notes on the assumptions that the authors make and try to challenge these assumptions. Ask yourself whether or not the authors are making reasonable assumptions, are they based on solid facts, and how these assumptions affect the soundness of the results. These questions can help you to understand the proposed solutions and the way in which the authors evaluate them.

Critically examine the experimental methods used in the evaluation. Did they (or can they) measure what they claim? If the paper is from the proceedings of the Very Large Database (VLDB) Conference, does it test against a petabyte-scale database? If the paper deals with distributed systems issues, are the nodes located on the same virtual machine host, or are they located in Australia, Asia, Europe, and the Americas? Can they explain their observations? Did they have adequate controls? Were tests carried out in a standard way?

Carefully examine the statistical analysis of numerical evaluation. Think about if the statistical methods are appropriate for the subject and if they are properly applied. Check if the author performed proper error analysis of experimental results before drawing a conclusion. Make statistical assessment of whether observations reflect a pattern rather than just chance to make sure the results is statically significant.

Approach the conclusions with scientific skepticism. Make sure that the result data supports the conclusions and the observations can infer to the conclusions logically. Think about if there are any other explanations for the observed results. Check the data to see if there are any correlations that are ignored by the author. By challenging the conclusion with all these questions, you can probably better understand how did the author has reached the conclusions and learn to apply the reasoning approach to your own research.

[1] Cited from [1]

Also make notes on what you don't understand, and put them right on the front page! Don't stop just because something is unclear or difficult; you will often find that if you push ahead, the authors will give you a clue that will explain the difficulty. Some researchers call this late-binding: holding content in their heads that they don't understand until, sometime later, "the penny drops". It's a useful skill to acquire. "Sometime later" may be weeks or months later, when some other paper gives you the insight that makes an earlier one "click". Once you understand, go back and add a note explaining the difficulty.

1.4.5 Goals of the fourth pass

The fourth pass is when you get down to detailed study; you go back over the hard parts or the parts that were unclear, and frame your own questions about what the authors leave unanswered. You may find that you do this pass on a minority of the articles that you read.

1.4.6 Remembering what you've read

When you are finished, write a summary of the article in your own words. If you have doubts about the authors' thesis, note your objections. If you find a phrase that you think you may wish to quote, note that too.

When you have finished reading a group of papers, write a summary of all of them. The summary should include the relationship between the papers, for example, how one paper's key idea is originated from another, or how the later article one improved something introduced earlier.

1.5 Cracking a tough paper

Some papers are particularly tough. Here are some strategies that can help.

- Have someone who understands the paper explains it to you.
- Try to explain what you find difficult to someone else.
- Find an article by the same authors targeting a more general audience.
- Make up (additional) examples. This can be a good way to master definitions, understand algorithms, or check out theorems. Working through an example is a good way to explain something quickly.
- Draw a picture.
- Write code.

- Recast portions in alternative notation, *e.g.*, logic, functional programming language.
- "Guess" what is going on. See if your hypothesis is consistent with the rest of the paper.
- Try to find a counterexample to a theorem. Often, this is the way that proofs are first found.
- Send questions to the authors. You may be surprised to find that authors are often happy to answer *specific* questions about their work. Phrase your question so that there is a 2–3 line answer; quote enough of the context so that the authors do not have to go back and read their original paper.

1.6 Organizing Source Materials

Over the course of your career, you are likely to read many hundreds, probably thousands, of technical articles. You will accumulate them steadily, and it's easy to become overwhelmed if you don't have a system for keeping track of them.

Any system that you use should be simple enough that you will keep on using it. It doesn't have to be perfect: if you can find the article that you are looking for within 2 or 3 minutes, your system is working fine. As an undergraduate, I simply filed research papers with the notes for the course to which they pertained. When I became a research student, there were no courses, so I made photocopies of the articles and filed them alphabetically *by first author* in a cabinet. Importantly, I also kept an index — on 3×5 inch cards — that categorized each article by keyword, sometimes several cards per article. I made it a rule that nothing went into the file cabinet without my first creating at least one index card. Other people group their papers by broad category, and then by author. Another commonly adopted rule is that whenever you make a copy of an article, *always* copy enough information to re-create a full citation. An easy way to do this is by copying the title page of the journal or proceedings.

I used to make paper copies of an article that I was working from, even if I had the full set of proceeding or the journal issue. That was because I could write freely on a copy, and because the copy was more portable. If you file your papers by topic, then filing is easier too. Today, a personal pdf file serves the same purpose as those old paper copies: you can annotate it, and you can carry it around.

Nowadays, I have replaced the file cabinet by a set of folders on my hard disk, one per author, but still organized alphabetically. These folders

currently contain about four thousand articles. The card index has been replaced by a *bibtex* database that contains the full citation information for each paper, as well as keywords of my choice that categorize it. Using *bibtex* for the index means that I can create the *References* section of an article with minimal effort; this used to be a major chore, and one that once caused me some embarrassment.

1.7 Chapter Summary

This chapter discussed techniques for effectively reading articles, and how these techniques might change depending on your motivation for reading. Here are the main ideas.

- Have your goal in mind when you read.
- Take notes as you read, and file them with the paper.
- Read in multiple passes; don't try to read from the front to the back in one session.
- Challenge the authors' assumptions and methods.
- Don't give up when the going gets tough.
- Don't be afraid to ask for help.
- Organize you technical articles so that you can find them in the future, and so that you can cite them correctly.

A handy one-page reference on reading is available at http://www.cs.pdx.edu/~black/ScholarshipSkills/LectureNotes/efficientReading.pdf. It summarizes the material in this chapter, and also contains some additional suggestions that you may find useful. It was prepared by Michael J. Hanson, and has been updated by Dylan McNamee and others.

Acknowledgments

This chapter is based on lecture notes originally prepared by David Maier, and revised over a period of many years by Todd Leen, Tim Sheard and myself. James Hook contributed the list on page 2, and Dylan McNamee the notes on reading.

Chapter 2

Background Research

If I have seen further it is by standing on the shoulders of Giants.
Isaac Newton

2.1 Introduction

Scientific discoveries are built upon prior work. Much of the motivation for publishing research is to allow other researchers to build upon it. But before they can build on it, they have to be able to find it! This chapter will help you to find prior work relevant to your interests.

In many fields, "doing research" is synonymous with undertaking a literature search. In computer science, there are many other forms of research, but the ability to do undertake a literature search is nevertheless a cornerstone of scholarship. Moreover, the state of the art in many areas of computing — especially the areas in which you are likely to be doing your own research — changes very quickly; this means that any published literature review that you find is quite likely to be out of date. Thus, being able to find more recent research on a given topic is a core competency for a computer scientist.

What is the output from background research? The results of a thorough literature search can themselves be valuable scholarship, and can be published as a "review" or "survey" article. Indeed, the ACM publishes a Journal — *Computing Surveys* — dedicated to such articles. Every doctoral dissertation contains a chapter that surveys the state of the field to which it contributes, and every published research article will contain a review of prior work, usually in a separate section. At the very least, your background research should produce notes for your own use, with references to

sources that you have read, more references to sources that look interesting but which you have not yet read, and annotations that explain the relationships of each of these references to each other and to your topic of interest.

2.2 Why Do Background Research?

The fundamental reason to undertake background research is to learn about an area of science, technology or engineering — perhaps because you have discovered that you need to *use* it in your own work, or because you plan to *contribute* to it, or sometimes just for curiosity. You will learn the concepts and terminology of the area, what other researchers consider to be the important issues, and what research techniques are employed — for example, simulation, proof, or statistical analysis. You will also learn what tools other researchers use to support those techniques. Perhaps as importantly, you will learn who the significant people are in the area, where they work, what conferences they attend, and where they publish their work.

If you are working towards a PhD, then at some point you will need to find a problem to work on. Although your advisor will probably help you with this, in the end the choice is yours, and you will need to know what sub-areas are active, and what other researchers consider to be the significant open questions. By the time you have finished your doctoral studies, you will know the relevant literature more thoroughly than your advisor. This knowledge will let you situate your own research in the context of the trends in the field, and help you to explain how your work differs from that of other researchers — which is an important part of a dissertation. There is nothing wrong with choosing to "buck the trend", and work in a sub-area or use a technique that others have discarded as unproductive. However, such a choice should be made consciously, and not out of ignorance! Finally, you need to be aware of others' work to avoid repeating it: every doctoral dissertation and most venues for the dissemination of research require novelty.

There are also more general reasons to undertake background research. You may just want to satisfy your cultural or technical curiosity, or to understand how contemporary scientists in a particular area operate: what sorts of experiments they conduct, how they analyze their results, and where they publish. What constitutes a "result" in machine learning or artificial life? How are these results obtained? How does research in computation differs from research in civil engineering or theoretical physics?

2.3 Why is Background Research Challenging?

As pointed out by Dr. Daveena Tauber, Professor of Research and Writing Instruction at Portland State University, conducting a literature review is arguably one of the most difficult parts of a research project. The reason is that reviewing the literature requires a high level of *meta-cognition*: the ability to think about thinking. In other words, reviewing the literature requires you to do more than understand the research works that you are reading: you also need to be able to discern the *relationships* between the works, to explore the reoccurring and exceptional patterns within them, and to find connections between the studied works and your own.

Doing a literature review necessarily requires a significant amount of reading. However, that reading process should not be passive: it should be an *active* process guided by a clear goal. There is likely to be a vast amount of literature in any active research field. Having a goal will let you decide which papers to skip after you have read the abstract, which you will skimmed for their main results and their bibliography, and which you will read in detail.

A literature search will usually cover many different kinds of publication, ranging from textbooks and monographs, through refereed journals and conferences, to informal research blogs. These kinds of publication have different characteristics. At one end of the spectrum, materials from informal research blog often provide innovative ideas, but comes without evaluation — and often with little but rhetorical justification. At the other end of the spectrum, materials like textbooks and published encyclopÃędias can be relied on as sources and are good for learning about terminology and fundamental principles, but they will not reflect the current trends in the field. Your search will need to draw on all of these sources, but each kind will need be treated appropriately, depending on the publication venue, the age, and the original goal of the source.

Finally, there is the problem of getting started. There is no single source that you can rely on: you will need to use a combination of approaches. Not everything of value is on the web, and not everything on the web is of value.

2.4 Where to Search

2.4.1 The Web

These days, most of us start our search on the World Wide Web, by typing a search term into a search engine like google or yahoo! Using precise search terms can help a lot. For example, if you want to know about the use of the *bootstrap* to estimate confidence intervals, you will have a lot more success if

you search on "statistical bootstrap" than if you just search on "bootstrap". The ranking algorithm used by your search engine can also influence what you look see. Google finds about fourteen million web pages in response to "statistical bootstrap", and Yahoo about two-and-one-half million. You clearly are not going to look at all of these results, but will confine yourself to the first page or two, and its quite likely that there will be a few results that show up in the first page of one engine that don't show up in the first page of another.

Wikipedia deserves a special mention. Wikipedia articles can be great places to start if you are looking for a quick overview of (or a refresher on) a basic topic. For example, I find that the mathematics articles are generally well written and often more accessible than my shelf of text books; they also frequently contain example, s graphs and animations that help me to understand a definition or a procedure. But the quality of Wikipedia articles varies wildly. A particular article may or may not have been written or reviewed by experts. Most articles will have a list of references, and these are essential reading if you decide that the topic is important and warrants further study. If there are no references, then the article itself may be nothing more than rumor, or the author's opinion.

In contrast, Wikipedia is not much help with active research topics. This is because the circle of people who know enough about the topic is small, and they generally have better things to do than write or edit Wikipedia articles. Wikipedia also has a guideline against publishing original research; this means that the research must first appear in a published and citable venue, and only later can be cited by Wikipedia.

2.4.2 Ask Around

Once you have gained some basic background from a Web search, and before you start to dive into the list of references for further reading that you created from the results of that search, I recommend most strongly that you *talk* to other people. Talk to teachers, researchers, colleges, friends, acquaintances met at conferences ... any won who might be able to suggest good places to continue your search. Don't be afraid to appear ignorant: instead be concerned with appearing eager to learn.

There are two reasons to talk, rather than to send email. First, the person to whom you are speaking will know that this is personal request, directed specifically at them. Second, it is far more economical of the other person's time. By asking you a few questions in return, they can quickly narrow down what it is that you want to know, and either suggest other people to ask, or direct you to an appropriate source.

2.4.3 Important skills for doing a good literature review

A good literature review is not merely an ensemble of summaries. Instead, those summaries should be connected to each other so that they tell a complete story about the research problem the authors are approaching. In her workshop "Write a Killer Literature Review", Dr. Daveena Tauber articulates the following core skills.

Discern: Reviewing the literature requires us to be active readers, as discussed above. Our sources will be from diverse venues and at varied levels; we need to keep our goal in mind as we read.

Synthesize: The more we engage with the materials that we read, the easier it will be for us to understand and absorb them. Have some questions or examples in mind: did Purfrock and Smee use the wobulation technique invented by Peacock and Simons two years earlier; can their method be applied to the example that Fredrick and Davidson introduced in the paper that I read this morning? Write your questions and answers down, in a place where you will find them next time you look at this article.

Evaluate: Evaluation is the critical part that most of the literature review ask for. This means that we should be able to identify, acknowledge and emphasize the merit and credibility of the works that we cover in our literature review. In addition, we need to be able to explain the contribution of each works, and explain their influence to other works in the field.

Narrate: This is perhaps the most important skill we need to have in order to make a good literature review. All the others parts can become useless if we cannot narrate or articulate all the information we have gained when we study the literature. In other words, to create an effective literature review, we need to be able to leverage the insight we gain from each works as well as the relationship between different works that we surveyed to connect them in such a way that they compose a complete story that relate to our research.

2.5 Conducting a Literature Search

This section describe the process of literature search. Back in the old days, search for literature involved going through hundreds of catalogs in the library. With modern computing power, there are powerful search tools and gigantic database that can help a researcher conduct the search task more effectively.

However, in order to conduct an effective search, you need to know how to do it right. Section 2.5.1 talks about developing the necessary background to conduct a literature search. The outcome of this is a set of effective search terms, or search goals. One we have these, Section 2.5.2 describes different search tools, and search strategies to conduct literature search.

2.5.1 Developing background knowledge

Before searching for anything, it is important to construct a set of questions that can guide your search in the right direction. Searching with a clear set of questions in mine will help you quickly build the big picture of the research you are going to conduct, the current challenges in the field, and the implications for future research. Here is an example of a bad and a good question.

Bad: How can I figure out by looking at a program whether it will stop or not, and can I make programs simpler so it's easy to see they will always stop.

Good: What is known about automatic termination analysis? Are there subrecursive languages that guarantee termination?

To be able to construct good search questions, you need a substantial background knowledge in the field. You may be tempted to just go to a search engine and dwell through millions of web page to "brute-force"âĂİ learning about the field. However, this is not a good idea because without good search queries, you may not get good search results, and the search task become ineffective and time consuming. You need to have a list of specific terminologies in the field in order to conduct effective search queries. The process of understanding the background knowledge of what you are researching will help build the terminologies collection.

It is helpful to develop background knowledge by skimming materials that can give you a big picture of the field. This process will allow you to get familiar with the field and build a set of effective terminologies for searching. Below I describe a variety of approaches to facilitate skimming. While doing so, please remember to write down all the terminologies, and make sure to keep this process short. You should not spend more than one week just to build up background knowledge.

Textbooks First of all, textbook is a good choice because it is written in ways that make it easy to understand. An up-to-date textbook may provide valuable discussion on the field, allow you to get familiar with terminologies, and explain concepts succinctly. You may have a problem finding the right textbook because most textbooks cover a wide range of topics within a field, while your research only focus on a very specific problems. It is often

not possible to find a textbook that only talks about your topic. So when searching for one, look for each chapter content to see whether they discuss the problem you are interested in.

Professional venues It is also helpful to write down a list of top publishing venues for the topic you are researching. These venues often provide valuable tutorials and survey papers on a variety of topics. These materials are like textbook, but focus only in the field you are interested to. Looking at the program can also let you know about the current trend and hot topics.

It is important to only look at highly ranked conferences, workshops, or journals. To find out whether a publishing venue is good or not, you can do a quick search for its ranking on the Internet. It would be much more efficient to just ask someone who is experienced in the field, like your advisor, or a professor in the department who is working on that field. You can also personally see if a conference is good or not by looking at the conference website. Look for information such as the number of participants, acceptance rate, and the list of sponsors.

Even if you cannot find any helpful readings at all, it is also useful to skim through the program. Nowadays, most conferences divide its program into specific sub-topics. You should try to locate the topic that is closest to yours, and then look at the list of papers and authors. It helps you to identify researchers or research group that are actively working on similar topics. You may be able to find dissertations that cover some parts of your field.

Ask around Another approach to quickly fill up your need for background knowledge is to ask around. Ask your advisor, other professors in the department, or other graduate students. People who are experienced in a topic can quickly give you a big picture in a short time. However, you should do the work yourself first before coming out and ask around. By educating yourself about the topic, you should be able to extract the most valuable information from a conversation, and not waste everybody's time. Also, don't be afraid to branch out. Try meeting with people from different departments, if your research spans on different disciplines. You can even visit another university, if they are nearby. When you do so, send an email ahead of time to schedule a visit. It is also helpful to see a problem from the industry perspective. Check for local tech events in your area for events that are relevant to your topics. You can find a lot of experts when attending these events. The website http://calagator.org/ updates daily on technical events in the Portland area.

2.5.2 Searching

The outcome of Section 2.5.1 should be a clear goal and a list of terms in the field. Now you are ready to mine the web for more serious materials. To search effectively, you need a good search tool. Below, I will first present several search tool to search for literature in Computer Science. In the next section, I will describe several search strategies.

Google Scholar Google Scholar http://scholar.google.com/ is a web service developed by Google, released to the public in 2004. This service indexes scholarly literature from many sources on the web. One of the main advantage of Google Scholar is that it indexes the content of the paper, and the search engine is backed up by the search power of Google. In this way, keyword searching on Google Scholar is highly effective. The service also allow you to customize and filter the search. For example, you can filter the search results to display only the results from a specific conference, or in a specific range of years.

A potential problem with Google Scholar is that it may turn up literature from other disciplines if your search term is not sufficiently specific.

CiteSeer CiteSeer http://citeseerx.ist.psu.edu/ is a search service made available in 1998. It is widely known as the search engine for academic papers in computer science. CiteSeer excels in its ability to allow the user to perform citation search. Citation search will weigh search results by how heavily cited they are. In this way, you can easily find highly cited papers.

Publishers' websites. Academic publishers such as ACM, IEEE, Springer, and Elsevier allow subscriber to find and download papers through their website. Some of these websites also give you the ability to refine your search results. In Computer Science, you can check out the following resources.

- ACM Digital Library http://dl.acm.org
- IEEE Xplore Digital Library http://ieeexplore.ieee.org/Xplore/home.jsp
- Springer Link http://link.springer.com/

DBLP http://www.informatik.uni-trier.de/~ley/db/ is a computer science bibliography website. Started in 1980, it collects more than 2.3 million papers; its coverage is 10% more than CiteSeer. DBLP excel in its very complete bibliography on database research.

Search strategies In this section, I will describe several search strategies.

Forward search

Backward search

2.6 Supporting Technologies

2.7 Chapter Summary

Acknowledgments

Chapter 3

Core Rules for Better Writing

If we are facing in the right direction,
all we have to do is keep on walking.

Zen Proverb

3.1 Introduction

Writing well is not a gift: it is a skill that can be learned. The words that we use in our writing are the same words with which we think. As we improve our writing skills, we also improve our ability to think — to organize arguments, to frame issues in compelling ways, to avoid irrelevant issues and to arrange the facts into a smooth and complete story. Improving your writing skills will help you to become a better researcher, as well as a more successful publisher.

In this chapter we will provide some core rules to help you improve your writing. They don't cover everything you will need to know, and they certainly don't try to capture all of the details that you will find in a writing style guide, but they will help you build a firm foundation on which you can develop other skills. If you are not already following these rules, they can make a dramatic difference to you writing.

3.2 What is Good Writing?

Present to inform, not to impress.
If you inform, you will *impress.*

Fred Brooks

Good writing is easy to read. A good paper educates the reader without frustrating him or her. As Fred Brooks indicates in the above quotation, we should write to be understood, not to impress. Use clear, simple words and sentence constructions; avoid fancy words and "academic" usage. Be concise: use just enough words to convey your meaning clearly, but don't waste the reader's time with extra words. Thus, the "golden rule" of writing is that, as an author, your job is to ease the job of the reader. If in doubt about how to choose between two way to express something, choose the one that makes the job of the reader effortless, even if it is more work for you.

Enlist readers to help you make this choice. The core difficulty of good scientific writing is that to explain something simply, you have to know a great deal about it. And once you know a great deal about it, it's hard to remember how ignorant you were, when you didn't know anything about it. But you must try, for your job is to write for the reader who does *not* already know.

3.3 The Core Rules of Writing

Words have to be crafted, not sprayed.
They need to be fitted together with infinite care.

Norman Cousin

The "Core Rules" in this section will help you write better. The rules do not cover everything that you need to know about writing, but they do provide guidance that, if followed, will dramatically improve your writing.

3.3.1 Use Active Verbs

It used to be traditional for scientific works to be written in passive voice. Because the passive voice makes it easy to avoid saying what caused an action, it was thought to somehow be more objective. This is no longer the case, certainly in computer science. Active voice is more specific, and therefore more informative and clearer than passive voice; it is also frequently more concise. Moreover, active voice is more direct and forceful. If you do nothing more than concentrate on using active voice, your writing will improve.

What is Passive Voice?

First, let's remind ourselves what passive voice is. Take a look at this passive voice sentence:

> The file was initialized by the experimenter.

Who is doing the action in this sentence? Clearly, it's "the experimenter", even though the grammatical subject of the sentence is "the file". However, "the file" is *passive*: it isn't *doing* anything. Phrasing the sentence in the active voice — "The experimenter initialized the file" — places the actor as the grammatical subject, and replaces the verb "to be" by the active verb "to initialize".

Train yourself to spot passive voice constructions. They are characterized by some form of the verb "to be" (*am, are, is, was, were, have been, will be, will have been* followed by a past participle — that form of the verb that typically ends in "–ed", such as "initialized", or "improved". (English being English, there are irregular verbs whose past participles don't end in "–ed": "break", for example, has "broken" as its part participle, not "breaked", and "thrust" stays as "thrust" and does not become "thrusted".)

Here are some examples that illustrate the benefits of active voice.

Active voice is more informative

In the passive voice, the subject is missing. In the active voice, the subject is explicit, so the sentence is more informative.

Passive voice:		*Active voice:*
It is felt that you should re-write your research proficiency paper.	→	Professor Black feels that you should rewrite your research proficiency paper.
To get advice, an expert must be consulted.	→	To get advice, you must consult an expert.
Speech features can be mapped using a probabilistic approach.	→	Stylianou et al. [22] map speech features probabilistically.

Active voice is more concise and more forceful

Passive voice:		*Active voice:*
The glass was broken by Tim.	→	Tim broke the glass.
My first visit to Boston will always be remembered by me.	→	I shall always remember my first visit to Boston.[1].
There were a great number of dead leaves lying on the ground.	→	Dead leaves covered the ground.

The reason that active voice is more concise and forceful is that in English, we expect verbs to be action words. Taking an action and turning it into a noun weakens the the structure of the sentence.

noun form:		*verb form:*
There was a prohibition against the export of Rubles.	→	The Russian government prohibited the export of Rubles.
Anna made a translation of the poem from Polish into English.	→	Anna translated the poem from Polish to English.
Josephson obtained confirmation of these results.	→	Josephson confirmed these results.

The first example is conventional passive voice. The second and third examples are not: "Anna made" and "Josephson obtained" are, grammatically, active constructions. However, the verb forms are more forceful, because in these forms the action that we are interested in — translation and confirmation — is represented by a verb rather than by a noun. They are also more concise, because the auxiliary verbs "made" and "obtained" have been banished.

3.3.2 Put Key Ideas in Lead Position

Putting key ideas in lead position helps to prepare the reader for what's coming, and also allows the reader to skim efficiently.

Use this technique in each section section of a paper, except the introduction. The first paragraph in a section should act as an introduction to that section and summarize its key ideas. The following paragraphs should provide the details of the key ideas. Thus the paragraphs within a section move from the general to the more specific, from the most important to the least important.

You should also use the technique of putting the key ideas in the lead position within each paragraph. The first sentence in a paragraph should explain the purpose of the paragraph; subsequent sentences explain or develop the topic in more detail. In general, sentences within a paragraph move from the general to the more specific.

However, there are some exceptions. The major exception is that the *trailing* positions also powerful: the conclusion of article, the last paragraph of a section, and the last sentence of a paragraph. Don't waste these positions by filling them with digressions or unimportant details; instead, use them to restate your key ideas, or to state an important consequence of these ideas.

Another Exception is when you're trying to persuade your audience to accept a radical idea. Then you can lead them along and give them the key idea at the end, like the punchline to a joke. This is frequently used in a mathematical development. Here is an example.

> Consider Newton's second law:
>
> $$F = m\frac{dv}{dt} \tag{3.1}$$
>
> and integrate both sides over a displacement x
>
> $$\int_0^x F dx = m\int_0^x \frac{dv}{dt}\, dx \tag{3.2}$$
>
> substitute ...
>
> Hence we have shown that work equals the change in kinetic energy.

However don't overuse this format, particularly when the argument is long. If you do use it for a mathematical development, it helps the reader if you use the lead sentence in the paragraph to tell them what you're about to do. That is, it helps to give away the punchline! Remember that, unlike the stand-up comedian, your job as a writer is not to appear clever: it's to help your audience to understand. Here is the same example, re-written in this style.

> Now we show that work equals the change in kinetic energy.
>
> Consider Newton's second law:
>
> $$F = m\frac{dv}{dt} \tag{3.3}$$
>
> and integrate both sides over a displacement x
>
> $$\int_0^x F dx = m\int_0^x \frac{dv}{dt}\, dx \tag{3.4}$$
>
> substitute ...

3.3.3 Don't make Unsubstantiated Statements

In a scientific paper, statements of belief or of fact must be backed up. It is not OK to simply claim that "Refactoring tools are underused": you must substantiate that statement. This can be done in one of three ways:

1. referring to a specific result of the current work, for example "In Section 3 we show that refactoring tools are underused", or

2. citing your own prior work, for example, “We showed in our 2009 ICSE paper [3] that refactoring tools are underused”, or
3. citing the literature, for example, “Dig and Delany showed that refactoring tools are underused [5]”

Don’t use phrases like “It is common knowledge . . . ” or “It is generally believed . . . ”, or “Several researchers have shown . . . ”. Instead use constructs like "Hartman [23], Goolickan [24], and Brotman [25] show that . . . " or “Recently several researchers have shown that . . . [23,24,25]” or “Many researchers (see for example, references 23, 24 and 25) believe that . . . ”.

The exception is for things that really *are* common knowledge. It’s OK to say that the sun rises in the East without giving a citation, but there is no need to say that “It is common knowledge that the sun rises in the East”. If you think that the reader might not know that this is common knowledge, then you should give a citation: saying that it is common knowledge is just one more unsubstantiated statement!

3.3.4 Be Concise

Avoid wordiness. Here are some examples:

Instead of these	Use these
due to the fact that	because
in light of the fact that	since
in order to	to
in view of	since
on behalf of	for
after this is accomplished	then
in case	if
along the lines of	like

3.3.5 Be simple

Avoid “fancy” words. Here are some examples:

Instead of these	Use these
subsequently	next
modification	change
necessitate	require
endeavor	try
demonstrate	show

utilize	use

3.3.6 Use a Consistent Lexical Set

Use the same word to refer to a concept throughout the paper. For example, don't use "swoondigger" in one place, and "pomponicator" in another place, unless they are generally accepted synonyms. Although it may be obvious to you that these are synonymous, it may not be obvious to the reader.

When you make an exception and use several terms to refer to the same concept, make sure you tell your reader that you are using the terms synonymously: "Here we use a swoondigger (also referred to as a pomponticator), to force coherence between ..."

3.3.7 Define terms when first used

Don't make the reader guess what you mean until the last section of the paper, where you finally get specific about the meaning of a term. There are two types of definitions:

- Full sentence: The term has more than one meaning, or the audience probably does not know the term, or the definition is controversial (no agreed standard definition).
- Phrases: The term has more than one meaning and you are clarifying which meaning you adopt. The audience probably does know the term, but you are including a definition to avoid ambiguity, perhaps because the definition was controversial in the past.

3.3.8 Avoid one-word sentences and single sentence paragraphs

Never use one-word sentences and try to avoid single sentence paragraphs. Single sentence paragraphs are usually an indication that there's a problem with organization. Figure out where the idea belongs. It may be that it's not important enough to develop into its own paragraph, but doesn't fit in one of the existing paragraphs. If that is the case, then remove it!

3.3.9 Re-write

All good writing is rewriting. When you re-write (especially a section, paragraph, or sentence that you know is difficult to read), keep asking yourself "what do I really mean?". Allow yourself several passes to get rough spots

really concise, really simple, and really clear. A good way to get more concise is to ask yourself if words in a sentence, or sentences in a paragraph are helping you make the key points, or whether they can be discarded.

3.4 Chapter Summary

This chapter discussed core rules for writing better articles. Pin them to your monitor so they are in front of you as you write! Here are the core rules.

- Use active verbs.
- Put key ideas in the lead position of sections and paragraphs.
- Don't make unsubstantiated statements.
- Be concise.
- Be simple.
- Use a consistent lexical set.
- Define terms when first used.
- Avoid single sentence paragraphs.
- Rewrite with an intent to make things simpler, more concise, and clearer.

Acknowledgments

The original draft of this chapter was composed by Kai Cong, based on lecture notes originally prepared by Todd Leen, and revised over a period of many years by Tim Sheard and Andrew Black.

Chapter 4

Mechanics

This chapter summarizes some mistakes that students and also more experienced academic authors commonly make in their papers. They are mostly related to grammar, punctuation and technical notation. They may seem trivial, but if you get them wrong you will distract, mislead, and slow down your readers.

Constructs that are grammatically correct may still not be optimal for clarity and concision.

4.1 Basic Grammar

The rules of grammar tell us how sentences are structured. English grammar is a lot more fluid than that of many languages, but there are still rules that should never be broken, and other that should be broken only with caution one you have given a lot of though to the alternatives.

4.1.1 Singular and Plural

Verbs should match the noun, not the adjective phrase's noun.

✗ **DON'T:** This set of guidelines are useful ...

✓ **DO:** None of the available programs works

Pronouns should match the original noun.

✗ **DON'T:** The operating system is the first bastion of defense against the hacker. Keep them up to date.

✓ **DO:** Operating systems are the first bastion of defense against the hacker. Keep them up to date.
Alternatively:

✓ **DO:** The operating system is the first bastion of defense against the hacker. Keep it up to date.

4.1.2 Past and Present Verbs

Be consistent throughout the paragraph except when the action is clearly in the past or future.

4.1.3 Articles

The choice between “A” and “An” depends on the sound, not the spelling, of the noun:

✓ **DO:** a university
, but

✓ **DO:** an honour
. This works for abbreviations too:

✓ **DO:** an MP
, but

✓ **DO:** a member of parliament
. Adjectives change the sound, so they change the article too:

✓ **DO:** a table
, but

✓ **DO:** an unusual table
.

'The' is used to indicate a specific object. Use 'a' or 'an' if talking about a generic object.

✗ **DON'T:** Tonight we are going out to see the movie.
If we are just going to see any movie:

✓ **DO:** Tonight we are going out to see a movie.
If we are going to see a specific movie, we can use "the movie" if it is clear what movie we are talking about OR we can be clearer by indicating which particular movie.

✓ **DO:** Tonight we are going to see the top movie of 2013.

4.2 Spelling

Use a spellchecker but don't take it for grant it that if the spellchecker indicates a word is okay, then it must be correct. All the spellchecker checks is whether a word is in the dictionary; not whether the word is correct in the context that it is in.

Some words have alternate spellings. Because of this, spelling should be consistent throughout. Also, use U.S. English spelling in U.S. conferences/journal papers.

4.2.1 Word Pair Confusables

Word pairs (or sometimes confusables) are pairs of words that are close in spelling but different in meaning so they are often confused by writers. It's good check these by eye as they are not picked up by spellcheckers (and might not be by grammar checkers either). **affect and effect**: Affect: is a verb meaning "to influence".

✗ **DON'T:** Background noise effected the quality of your results.

✓ **DO:** The network load affected the available bandwidth.

Effect: is a noun meaning "a result" or "bringing about a change in something".

✗ **DON'T:** The network load had no affect on bandwidth.

✓ **DO:** The network load had no effect on bandwidth.

Sometimes "effect" is a verb meaning "to cause something to occur," as in, "The cheerful music effected a change in Joe's bad mood." However, this usage is rare. It might be best to avoid using it as a verb unless nothing else works.

✗ **DON'T:** The cheery music effected a change on Joe's bad mood.

✓ **DO:** The cheery music caused a change on Joe's bad mood.

✓ **DO:** The cheery music led to a change on Joe's bad mood.

Likewise, "affect" is also a noun, but rarely used as such.

✗ **DON'T:** Joe had a sour affect. (Note: this is correct but confusing)

✓ **DO:** Joe was in a bad mood.

stationary and stationery: Stationary: An adjective meaning not moving

or immovable.

✗ **DON'T:** As the name implies, "base stations" are stationery.

✓ **DO:** The lighthouse was stationary because it stood on top of a rock.
Stationery: Paper to write on.

✓ **DO:** The love letter was written on perfumed stationery.

✗ **DON'T:** You can use LaTeX to print your own stationary.
complement and compliment:

Complement is defined as "goes well with" or "pairs with" something. Compliment is defined as 'praise'.

✓ **DO:** The white wine complemented the halibut steak.

✗ **DON'T:** The TA and professor complimented each other's skills quite well.

✓ **DO:** I complimented the professor on her choice of examples.

alternate *vs.* alternative: alternate: Switching between two things.

✓ **DO:** The current alternates between positive and negative.
alternative: Another choice besides the one given.

✓ **DO:** Vanilla was one alternative to chocolate.

imply *vs.* infer: see Dupré [2, §64]: imply: To put an alternative meaning into what is being spoken/written.

✓ **DO:** When you said "You took the money.", were *you* implying that I was a thief?
infer: To conclude that someone's speech/writing has an alternative meaning.

✓ **DO:** When you said "You took the money.", *I* inferred that you were accusing me of theft.
Notice that the *speaker implies* while the *listener infers*.

deduce, induce and infer: *ibid*

4.2.2 Counting and Mass Nouns

There are nouns that indicate a 'mass' of indistinguishable parts. For example, Water is a mass noun:

✓ **DO:** How much water do we have?

✗ **DON'T:** Too many waters filled the tub and it overflowed.

People are distinguishable and therefore can be counted. So 'people' is a counting noun:

✓ **DO:** Many people filled the square.

✗ **DON'T:** Less people filled the square than before.

Money is also a mass noun:

✓ **DO:** I had less money than my friend.

✗ **DON'T:** I had fewer money than my friend.

Two vs. Three or more

Use 'er' when comparing two things only. Use 'est' for comparing more than two things.

✓ **DO:** Mergesort is faster than Quicksort

, but

✗ **DON'T:** We tested six algorithms, Bubblesort was consistently the slower.

✓ **DO:** We tested six algorithms, Bubblesort was consistently the slowest.

Choose *between* two alternatives, but *among* three or more possibilities.

✗ **DON'T:** I had difficulty choosing among the two different paint colors.

✓ **DO:** I had difficulty choosing between the two different paint colors.

✗ **DON'T:** Between all six flavors, I like chocolate the best.

✗ **DON'T:** Among all six flavors, I like chocolate the best.

4.2.3 Parallel Clauses

4.3 Clear Writing

4.3.1 Word Choice

Term Consistency

'Like' vs 'Such As'

See Dupré [2, §27].

✓ **DO:** I'm looking for a shrub like a Hydrangea, but with evergreen leaves.

✗ **DON'T:** To prepare our papers, we considered using a text processor like LaTeX.

"Which" vs "That"

"Which" and "that" are not interchangeable. "Which" is normally proceeded by a comma, and introduces a short *description* of the thing in question.

✓ **DO:** The lawnmower, which is in the garage, is broken.
This sentence would mean the same thing without the parenthetical phrase ", which is in the garage," — but it might take you longer to locate the broken lawnmower.

In contrast, "that" is not preceded by a comma; the phase introduced by "that" *defines* the thing in question.

✓ **DO:** The lawnmower that is in the garage is broken.
Here the information contained in the defininig phrase "that is in the garage" is essential; the implicaiton is that there is more than one lawnmower, and the one in the shed works just fine.

✓ **DO:** The dinner, which was prepared by an expert chef, began at 9:00pm.

✓ **DO:** The dinner that was prepared by an expert chef began at 9:00pm.
Again, the implication behind the second sentence is that there were other dinners, and that the amateur chefs were able to serve their meals in a more timely fashion. See Dupré [2, §17], and Fowler [3, "that, rel. pron. 1"]

4.3.2 Word Placement

Word placement is critical to writing clear English. Because English is not an inflected language, the reader or listener is left using word placement as a clue to what the reader means. Poor placement leads to ambiguous statements, and can leave the audience confused. Some examples of word placement rules follow.

Placement of "above" and "below"

Above and *below* follow the word that they describe:

✗ **DON'T:** The below item is frequently ignored, to the detriment of clarity.

✓ **DO:** The item below is frequently ignored, to the detriment of clarity.
Dupré [2, §48] argues that we should instead use more precise terms.

✓ **DO:** Computer scientists frequently ignore the following item, to the detriment of clarity.

Placement of "only"

Only qualifies the word that immediately follows.

✗ **DON'T:** You took out the garbage this morning only.
All of the following are correct — but note how they differ in meaning!

✓ **DO:** *Only* you took out the garbage this morning.

✓ **DO:** You *only* took out the garbage this morning.

✓ **DO:** You took out *only* the garbage this morning.

✓ **DO:** You took out the garbage *only* this morning.
For further discussion, see Dupré [2, §5]. For another istrautive example, consider the possible placements of *only* in the phrase: "Yesterday I hit the man in the eye."

Split Infinitives

An infinitive is a verb form that can act as a subject or object, as in "to be" or "to laugh." A split infinitive occurs when the speaker places another word, typically an adverb, between "to" and the verb. This is considered poor grammar.

✗ **DON'T:** It is not good policy to incessantly criticize your students.

✗ **DON'T:** To boldly go where no adverb has been before.

✓ **DO:** It is good policy to provide students with appropriate feedback, both negative and positive.

✓ **DO:** It is good to laugh.
For further discussion, see Dupré [2, §38] and Fowler "Out of the frying pan".

4.3.3 Non-referential Pronouns

Non-referential pronouns, such as 'this' 'that' or 'it', can refer back to any noun. The meaning of one of these pronouns may be clear to the writer at the time when they write, but they may leave the reader confused. For example:

✗ **DON'T:** Reducing the number of service queues increases average delay and reduces the number of idle periods. This affects the recovery subsystem.
One can clear up the confusion with more specific language, such as:

✓ **DO:** Reducing the number of service queues increases average delay and reduces the number of idle periods. The delay and the reduction in idle periods affect the recovery subsystem.

4.3.4 Lists

Lists are assumed to be complete, unless the author indicates otherwise.

✗ **DON'T:** Programming environments — Eclipse, XCode, BlueJ — improve programmer productivity.

✓ **DO:** Programming environments — Eclipse, XCode, BlueJ, *etc.* — improve programmer productivity.

✓ **DO:** Examples of programming environments are Eclipse, Xcode and BlueJ.
Note that one only needs to indicate that the list is incomplete once; additional indications are redundant and can be removed.

✗ **DON'T:** Examples of programming environments are Eclipse, Xcode, BlueJ, *etc.*

4.3.5 Use of Jargon

Jargon is endemic in technical communication, and so it can be difficult to realize when one is using jargon in academic writing. Nonetheless, one must take care to limit use of jargon, because jargon can confuse outsiders. Consider having an outsider read the document and point out jargon.

✗ **DON'T:** "We had a high-bandwidth conversation with our users."

✗ **DON'T:** "Our survey showed most administrators bounced the server to resolve the problem."

✓ **DO:** "We had an intensive, probing conversation with our users."

✓ **DO:** "Our survey showed most administrators power-cycled the server to resolve the problem."

4.3.6 Use of Latin

Latin makes one's writing harder to understand, but there are rare exceptions where a Latin phrase is acceptable. The phrase "*et cetera*" (or "*etc.*") is still in common use even outside academia, and can be treated as English. The phrase "*et al.*" is acceptable because it's commonly used in academic literature, but one may also elect to use an English equivalent such as "and colleagues." Whether one chooses to use "*et al.*" or "and colleagues," one should consistently continue to use that phrasing through the entire work, and in one's bibliography.

✗ **DON'T:** Functional languages, *e.g.* Scheme and OCAML, often find use in artificial intelligence programming.

✓ **DO:** Functional languages, *such as* Scheme and OCAML, often find use in artificial intelligence programming.

✗ **DON'T:** *NB:* we still haven't shown that the smart card's cryptographic algorithm is in fact linear.

✓ **DO:** *Note:* we still haven't shown that the smart card's cryptographic algorithm is in fact linear.

4.4 Style and Notation

4.4.1 Citations

Writers often treat citations as nouns. In fact, they are not nouns; they are parenthetical statements, and should be treated as such. One can check for this by reading one's writing aloud.

✗ **DON'T:** In [dunmore1970], we see the first use of proof by nonexistent reference.

✓ **DO:** Dunmore [dunmore1970] was the first to call this technique "proof by nonexistent reference."

4.4.2 Numbers

See Dupré §24 and 34. Writers normally spell out whole numbers less than 10, but use numerals for units of measure, time, dates, page numbers, chapter numbers, percentages, money, proportions, part of a series of larger numbers. But strive for clarity:

✗ **DON'T:** 20 100-Mbit Ethernet ports.

✓ **DO:** twenty 100-Mbit Ethernet ports.

4.4.3 Decimal points

Use decimal points to express a level of precision in numeric values. Consider for example the difference between:

> "The rod was one-half inch in diameter."

and:

> "The rod was 0.50 inches in diameter".

The former expresses that the rod is "about" half an inch in diameter. The latter says that the rod is somewhere between 0.495 inches and 0.505 inches in diameter — it is more exact in its meaning. Neither of these is wrong; which is better depends on how precise the figure needs to be.

4.4.4 Capitalization

4.4.5 Fonts

4.5 Punctuation

Good punctuation is essential to clear writing. Bad punctuation, on the other hand, is distracting at best and confusing at worst. Consider for example:

✗ **DON'T:** For dinner we had chicken and Robert and Nancy came over.

This could mean:

✗ **DON'T:** For dinner we had chicken and Robert, and Nancy came over.

Or it could mean:

✓ **DO:** For dinner we had chicken, and Robert and Nancy came over.

Here are some guidelines for good punctuation.

4.5.1 Quotations

Quotations should either be in quotes or in a different font. In the USA, a printers' convention holds that commas and periods that logically follow a quotation are moved inside the closing quotation mark. (British usage is to punctuate according to sense.) This is problematic when you are quoting computer input.

✗ **DON'T:** Type "rm *," then press "return"
.

4.5.2 Colons and Semicolons

4.5.3 Commas

4.5.4 Solidus and Dashes

The solidus (/) means "per" or "divided by." Don't use it in text.

✗ **DON'T:** Input/output operations are time-consuming.

✗ **DON'T:** His/her productivity should be improved if he/she adopts the new tool.

See Dupré [2, §117].

You will probably use four different dashes: mathematical minus $(x - y)$, the hyphen, the en-dash [2, pp 218–221], and the em-dash — which I think looks best with a hairline space on either side. Know how to type each in your chosen tool!

4.5.5 Parenthesis

4.6 Further Reading

These examples are based on a set of PowerPoint slides called *l05Mechanics*. Read over the sides and consider adding to this document. http://www.cs.pdx.ued/~black/ScholarshipSkills/LectureNotes/l05Mechanics.pdf

Acknowledgments

Many of the examples in this chapter have been contributed by David Maier and Andrew Black. Others have been inspired by Lyn Dupré [2] and H.W. Fowler[3].

Chapter 5

Citing your Sources

5.1 Introduction

It is important to properly and appropriately cite the literature in scientific writing. This is for two reasons. First, as we explained in Chapter 3, one of the core rules of writing is never to make unsubstantiated statements. The way that we substantiate our statements is either to refer to another part of the current work, or to the scientific literature. A reader who doubts the veracity of a statement of fact should never be left thinking "I wonder where the author got that idea!" It should always be clear where "that idea" comes from, which means that we give the reader enough information to find the referenced work, and, in the case of longer works, the particular section or page that we used as our source.

One kind of research writing that is particularly rich in literature citations is the "literature survey" or "related work" section of a research paper, or the whole of a survey paper. Here, citations to appropriate sources show that you've done your homework and are aware of the background and context into which your work fits. You may also wish to include citations for sources that add relevant information to your work, or that present alternate views. Throughout your writing, citations help to validate your arguments, and provide avenues for interested readers to follow up on aspects of your work. In other words, they help to weave the web of science.

The modern way of citing the literature is to include a short *callout* in the running text whenever you need to refer to a source. Callouts are also known as citation *labels*. The label is treated as a parenthetical remark, not a noun, and like all parenthetical remarks, the text must make sense without it. The bibliography information for each of your sources appears once, in a list at the end of the document (or sometimes, in a book, at the end of a chapter). The term *citation style* refers to the way that the labels are composed and

formatted, and the way that the bibliography is formatted and abbreviated.

The citation style used in a journal article is usually dictated by the journal. In a conference paper, or a monograph, or a term paper, you have more freedom. Two basic styles of callout are in common use, although there are others. The first is the "Numeric" style, in which citation labels are simple numbers, usually set off from the text in brackets thus [8], but which in some publications are parenthesized (8), or occasionally superscripted[8]. Brackets are preferred, particularly in mathematical work, to avoid confusion with equation numbers and powers.

The second is the "Author–Date" style, in which citation labels take the form of the name of the author of the cited work followed by the year of publication, like this [van Leunen, 1979]. If the author's name has already appeared as part of the text, it is not repeated in the callout.

> A comprehensive reference for citation styles is van Leunen's "Handbook for Scholars" [1979].

5.2 When to Cite References in Scientific Research Papers

You should acknowledge a source any time — and every time — you use a fact or an idea that you obtained from that source. Thus, you clearly need to cite sources for all direct quotations. But you also need to cite sources from which you paraphrase or summarize facts or ideas: whether or not you've put the fact or idea into your own words, the fact or idea came from somebody else and you need to give them proper acknowledgement. The only exception is an idea that is "common knowledge," but if you didn't know that was common knowledge until you found it in a particular source, you have just demonstrated to yourself that it wasn't common knowledge, and you should cite an authoritative source, such as a textbook or reference book.

Sources that need to be acknowledged are not limited to books, conference papers, dissertations and journal articles, but include websites, computer software, written and e-mail correspondence, and even conversations with other people (in person or by telephone). Furthermore, if you use figures, illustrations, or graphical material, either directly or in modified form, that you did not yourself create or design, you need to acknowledge the sources of those figures. (In a formal publication, if the material is copyright, you *also* need obtain permission for reuse from the copyright holder.)

5.3 Citing References in your Text

When you cite a reference in your text you should normally use one of the following three forms.

1. Cite the publication, and then say what relevant facts or ideas you found there. For example:

   Numeric Style: According to Rodgers [6], the Appalachian mountains were formed in three events.

   Author-date style: According to Rodgers [1983], the Appalachian mountains were formed in three events.

   Mention the author by last name in the sentence, and follow that with a non-breaking space and the citation callout. In author–year style, the author's name should not be repeated in the callout.

2. First give the facts or ideas mentioned by the author, and then attribute these facts or ideas by citing the source. For example:

   Numeric Style: The first of the three events occurred in the Ordovician, the second in the Devonian, and the third in the Carboniferous and Permian Periods [6].

   Author-date style: The first of the three events occurred in the Ordovician, the second in the Devonian, and the third in the Carboniferous and Permian Periods [Rodgers, 1983].

3. Quote the author exactly; be sure to put the quoted phrase between quotation marks, or set it off typographically. Then give the citation.

   Numeric Style: "All the climaxes produced mountainous islands or highlands that shed vast amounts of debris westward to form clastic wedges or delta complexes on the continental margin." [6, p. 229].

   Author-date style: "All the climaxes produced mountainous islands or highlands that shed vast amounts of debris westward to form clastic wedges or delta complexes on the continental margin." [Rodgers, 1983, p. 229].

You need to include the page number in the citation if you are quoting directly, or if the source is very long and the specific fact or idea you are citing can be found only on a specific page. Direct quotations that are 4 lines long or more should be set off from the rest of your paper by use of narrower margins and single spaced lines; shorter quotations can be run into your

paragraph, and enclosed in quotation marks. Your goal is to make it quite unambiguous which words come from the cited source, and which are your own.

If you are using author–year format, and have more than one source by the same author published in the same year, distinguish them both in the citation and in the reference list by appending the letters *a*, *b*, *c*, ... to the year, in the order in which the citations appear in your paper. (For example: Allen 1996a, 1996b.)

If the reference you are citing has two authors, use the following format:

> Periods of glaciation have a large effect on sea level [Ingmanson and Wallace, 1985].

If the reference you are citing has more than two authors, use the following format:

> Hot spots are formed by the drift of plates over mantle plumes [Vink *et al.*, 1985].

Personal communications are generally not included in the list of references, so the citation cannot take the form of a short label that indexes that list. Instead, you will need to give more information in the citation itself. If your source of information is from a personal oral communication, you would use the following format for the first citation from that person:

> It is possible to correct the raw dD values measured on the mass spectrometer [Mark Conrad, Lawrence-Berkeley National Lab, personal communication].

Later citations to the same person can be shortened, as in:

> The reproducibility of dD determined by these methods is thought to be about +/- 2 per mil [Conrad, personal communication].

If your source of information is from written correspondence (a letter or e-mail), substitute the word "written" for the word "personal" above, and add the date of the letter or email. Unpublished papers, technical reports or manuscripts that readers might be able to locate for themselves should be listed with the published references.

If your source of information has no individual identifiable author, use the name of the organization to which the work can be attributed in place of the author's name:

> The Java Virtual Machine Tool interface is a replacement for two older interfaces specific to profilers and debuggers [Oracle, 2011].

5.4 Formatting the Reference List

Your list of References should include all of the references you cited in your paper, and no more! It is normally arranged in alphabetical order by the last name of the first author. (A few publications that use the numeric style instead require that the list be sorted in the order that the citations appear in the text.) If you have more than one entry by the same author, they should be sorted by increasing publication date, so that the more recent publications come last. If you have multiple sources from a single author published in the same year, distinguish them both in the in-text citation and in the reference list, by appending the letters *a*, *b*, *c*, ... to the year, in the order in which the citations appear in your paper.

The point of the reference list is to include enough information for your readers to be able to find these sources on their own. The exact format is not critical, but consistency and completeness is. Reference lists are generally reverse-indented, that is, the citation labels (if you are using numeric format) or the author surnames (if you are using author–year format) are in the margin. This helps the reader to find references corresponding to specific citation labels that much faster. Follow the examples given below and you will be all set.

5.4.1 Books

List all authors by last name and initials, separated by commas if there are more than two authors. Put the word "and" before the last author in the list. Then put the year of publication, the title of the book (in italics), the publisher, the city, and the number of pages in the book. One author:

van Leunen, M.-C., 1992. *A Handbook for Scholars*, Oxford University Press, xi + 348 p.

Two or more authors:

Ingmanson, D. E. and Wallace, W. J., 1985. *Oceanography: An Introduction*, Wadsworth, Belmont, CA, 530 p.

5.4.2 For Articles from a Compilation

A "compilation" is a book in which each article or chapter has its own list of authors. The compilation as a whole will have an editor. List the author(s) of the article using the same format given above for books, then give the year, the title of the article or chapter (no quotes, italics or underlines). Next give the name(s) of the editor(s) of the book or compilation, followed by "ed." or

"eds.". Then put the title of the book (in italics), the publisher, the city, and the page numbers where the article can be found. Page ranges are written with an *en-dash*, which is longer than a minus sign but shorter than the dash used for punctuation (called an *em-dash*).

Rodgers, J., 1983, The life history of a mountain range — Appalachians, in Hsu, K. J., ed., *Mountain Building Processes*, Academic Press, Orlando, p. 229–243.

5.4.3 For an Article from a Journal or Magazine

List the author(s) of the article using the same format given above for books, then give the year, the title of the article or chapter (no quotes, italics or underlines), then the title of the journal or magazine (in italics), the volume of the journal (in bold face) and the issue (in italics), and page numbers where the article can be found. It is conventional *not* to use the month of publication, and even the issue number is sometimes omitted, provided that the page numbering of the volume is continuous.

One author:

Maddox, J., 1987. The great ozone controversy, *Nature*, **329**, p. 101.

Two or more authors:

Vink, G. E., Morgan, W. J., and Vogt, P. R., 1985. The Earth's hot spots, *Scientific American* **252**(*4*), p. 32–39.

5.4.4 Internet sources

Give the author's last name and initials (if known) and the date of publication (or most recent modification). Next, list the full title of the work (*e.g.*, the specific web page), and then the title of the complete work or site (if applicable) (in italics). Include any version or file numbers, enclosed in parentheses. Most important, provide the full URL to the resource, including the protocol. Be sure to spell this out correctly; it's best to use "copy and paste" for this. Finally, specify the date on which you most recently accessed the site, enclosed in parentheses.

Focazio, M.J., Welch, A.H., Watkins, S.A., Helsel, D.R., and Horn, M.A., 1999. A retrospective analysis on the occurrence of arsenic in groundwater resources of the United States and limitations in drinking-water-supply characterizations, U.S. Geological Survey Water-Resources Investigation Report 99-4279, http://co.water.usgs.gov/trace/pubs/wrir-99-4279/ (Accessed 16th October 2012).

Adapt these formats as necessary for other kinds of sources, including unpublished reports or manuscripts. The key idea is to be sure to include sufficient information for your readers to be able to find these sources themselves, without taking up more space than necessary.

5.5 Automating the Bibliography

With the advent of computerized typesetting, the process of constructing a bibliography for anything more than a short paper is invariably automated. The process works like this.

1. The bibliographic information for all of the referenced sources is collected in a database. Most publishers provide bibliographic information for their publications online, so job of the author is to collect these, correct them (yes, they are often buggy) and build a database file. I have a single database file for every technical publication that I have ever read that I think that I might, someday, want to cite; other authors construct a separate database for each article they write.

2. Each entry in the database has a unique key. The key is usually made up of a prefix of the first author's name and publication date (for example, *ingman1985*, but it can be any unique string (for example, *HandbookS*).

3. When you need to cite a source in your text, you insert a special text string containing the unique key of the source. This is called a *raw callout*. The exact form of the raw callout depends on the bibliography tool that you will use in the next step. Using BIBTEX, I would insert `\citep{ingman1985}`; other tools use delimiters like { and } or [. and .].

4. The whole document is processed by the bibliography software, which locates all of the raw callouts, extracts the bibliography information for the cited sources from the database, formats this information into bibliography entries, sorts them, arranges them at the end of the article, and calculates the citation labels (the numbers, or the author–year strings) that will be used in the finished document. Then the software takes another pass over the text of the document, and replaces the raw callouts with the citation labels.

5. The whole document is then formatted for a final time.

Some bibliography tools connect to particular document formatting software, and use the cross-referencing capabilities of that software to hide the raw citation keys and always display the citation in its finished form. This is convenient, but ties you to always using that particular combination of

bibliography tool and formatting software. If you find, partway through a project, that they aren't up to the job, you are stuck with a lot of re-work.

Maintaining a bibliography database takes some effort, but this effort is distributed over the whole of one's life as a writer. The payback comes when one wants to write an article with scores of citations: the bibliography is produced automatically, in seconds. Done by hand, this is a chore that used to take many hours. Moreover, numeric citation labels will be automatically adjusted when you find the need to add an additional citation.

An additional benefit of using a bibliography tool is that each publisher, and often each journal, has its own style for both citations and reference lists, and insists that its style is followed slavishly. Most bibliography software comes with an extensive library of styles, often numbering into the hundreds. Changing the style of the finished document can then be accomplished by changing a line or two of the input; without a tool, reformatting the bibliography would be a major project.

In this chapter, the numeric citations and the list of references were produced using BIBTEX and the *natbib* package. The *natbib* package can be used to produce either author–year or numeric citations, depending on the option that you select. This book uses the LATEX commands `\usepackage[square,comma,numbers,sort,sectionbib]{natbib}` and `\bibliographystyle{plain}`. Replacing *numbers* by *authoryear* will produce author–year labels. The examples of references in the body of this chapter were set by hand. You will probably see some small differences in style between them.

Acknowledgments

The content of this chapter is based on a web page compiled by Timothy T. Allen, as revised in 2000, which in turn expanded upon a handout originally prepared by an unknown author for distribution to students in introductory earth science courses at Dartmouth College. It has been updated and revised with contemporary information about computer science citation styles by Andrew Black.

Chapter 6

Presentations

I hate quotations.
Tell me what you know.

Ralph Waldo Emerson, *Journals*, 1849

6.1 Introduction

Suppose that you need to find out about a new topic. What do you do? You can of course find a good book on the subject, or a technical paper if the topic is too specialized for a book. You can go to the internet, and find information about your topic that ranges from a high-level overview to detailed research reports. However, if you have the opportunity, the fastest and most effective way to learn about a new topic is to *talk to someone* about it. Interestingly, that "someone" need not be an expert in the topic, although they do need to know a little more than you do. Surprisingly, it's often more useful to talk to a novice; that is, someone who has just begun learning about the topic, than to talk to an expert.

Assuming that you agree with the claim that talking is more effective than reading — and this is a generalization that, while true for most people most of the time is not necessarily true for all people all the time — *why* is this so?

In a conversation:

- the content can be tailored to the audience;
- the parties can communicate non-verbally — with looks and gestures — to indicate that a point is simple, and the speaker should move

on, or that something is confusing, and needs more discussion;

- the listener can ask questions;
- the parties can draw a picture and point to it, to help focus the discussion or to communicate a difficult point.

Why does a novice sometimes make a better teacher than an expert?

- The expert may have forgotten what it was like not to know; the novice can remember that only a few days or weeks ago he or she was equally puzzled by just that point.
- The expert may use technical jargon, forgetting that most people don't know what it means — or even that the terms in question *are* jargon.

It is precisely because talking is more effective than reading that *every* academic and professional meeting includes presentations. At academic conferences, being scheduled for a presentation normally requires that one first submit a paper to a panel of referees, and, if it is accepted, then revise the paper to meet their concerns. The revised papers are then published, so in a sense the presentations are superfluous: attendees could just read the papers. However, in practice, the presentations — at least, the good ones — are the focus of the meeting. A good presentation is not "reading a paper". Instead, a good presentation approximates what goes on in a personal technical conversation; the more we can make the presentation like a conversation, the more effective it will be.

6.2 Preliminaries

When you meet someone to begin a conversation, you don't plunge right into a technical topic: you first exchange some pleasantries, ask about their journey or their family, and generally get to know them and their state of mind. Then you establish the parameters of the conversation: how much time is available, what topics will be covered, and whether this will be an isolated conversation or part of a series.

A good presentation starts in the same way. The first priority is to get to know your audience. If you are presenting to a seminar group of which you are a member, you will know the audience before you start. If you are visiting another institution, you should ask your host about the likely audience before you arrive. Will the audience be academics from your specialty, or from another discipline? Will it include graduate and undergraduate students? With what background? If you are speaking at a conference or workshop, find out who typically attends. For example, if I'm talking about my

research in programming tools, I will give a different talk at a conference on human–computer interaction — where I would presume the audience to know about designing and evaluating interactive tools — than I would at a conference on software engineering — where I would presume the audience to be familiar with the process of software development.

Why is it important to know what your audience already knows? Talks are invariably time-limited, so you will always be faced with having to decide what to leave out. If there is something that you can rely on every member of the audience already knowing, that's one thing that you don't have to say. Nevertheless, be careful not to assume that the audience knows a piece of background information that is vital to your whole talk without checking first.

6.3 Kinds of Talk

There are various kinds of talk, characterized by their length and by the size of the audience, which governs their formality.

Standard Conference Talks. Most conference presentations are 15–30 minutes; the audience may be as small as 50, but has no upper limit. A listener can take away at most one or two ideas. Treat such a talk as an advertisement for your research — your goal is to persuade listeners to look further, in particular, to read your paper.

Long Conference Talks (invited talk, special slot). An hour-long talk to a large audience (100 people or more). The goal of such a talk is to educate the audience on the research. Listeners can take away at two or three ideas.

Class presentation (usually: 10–15 minutes). These presentations are directed at your peers and the instructor. Your goal is to demonstrate mastery of the class material. You can take advantage the common background and experience of the audience. It is similar in many respects to a conference talk, but will usually be geared to an audience or 50 or less.

For all kinds of presentations, gathering material for and creating good visual aids is a good starting point..

6.4 Gather Material

Gathering supporting material for your slides is an important task in the preparation step of your presentation. Before finding material, create a

rough list of ideas that you want to present in your talk. Then, knowing how long your talk will be, figure out how much material you will need to illustrate the points on your list.

6.4.1 How much material do you need?

There are two approaches to collecting material: too much, and too little. Both can work.

Some presenters try to gather three times as much material as they think they will need. This works for them because the process of gathering material is also an aid to understanding more deeply what they will present. Having a plentiful supply of material helps them pick the best examples and illustrations for the talk. Gathering more material also builds up their background on the topic, which boosts their confidence and will help them to answer questions that go beyond what they eventually present.

Other presenters will start with a bare outline, and fill it in with examples and illustrations until they have a talk of the appropriate length. They never have more material than they need, so they don't waste time collecting material that they don't use. This approach can work well if you already have a good understanding of the topic, and know where to look for supporting material.

6.4.2 Where do you find material?

If you are presenting your own work, for example, at a conference, the primary source for your material will probably be your own accepted paper, but you may also have results or experiences that post-date the paper. You should certainly include such material if you have it.

If the talk is an overview of work primarily done by others, such as an invited survey talk or a class presentation, you will need to inform yourself about the topic before you can inform the audience. Fortunately, there are many of sources for supporting material: relevant research and survey papers, books, tutorials, webpages, and videos. Cite your sources on your slides. If you borrow from other people's slides, acknowledge them, and make sure that you understand what the slides are saying.

6.4.3 Kinds of Material[1]

In general, there are four types of commonly used supporting materials: statistics, definitions, examples, and comparisons and contrasts.

[1] This part is based on the book "Business Communication" by Madhukany Jha

Statistics

Numbers are one of the most useful kinds of evidence that something needs to be improved or that one thing is better than another. Numbers are more powerful than words when you show the performance of a system or algorithm.

Definitions

Sometimes you will need to explain the terms you use in your presentation. The way that you use definition depends on who you are talking to. If most of your audience are experts in your topic, then use the accepted formal definitions. But if most of your audiences are not in your research field, you may need to rewrite the definitions, or use graphics instead of words.

Examples

A good example can save you a lot of time explaining a complex idea, and can liven-up your presentation. Read your examples carefully, and find the best examples to support your talk; you may find that you now know more than when you wrote the paper, and can construct a better example. You may also find that you can replace a formal definition with an example.

Comparisons and contrasts

When you need to explain something new to the audience, its often a good idea to compare it with something with which the audience is familiar. For example, you might describe a new machine learning algorithm like "Adaboost" by comparing it with its earlier version "Boost".

6.5 Create the slides

Slides should be used as illustrations for your talk; they give you pictures, diagrams, data and examples that help highlight your key ideas. Slides should be created to complement your talk, and not as a substitute for it.

6.5.1 Title and Introduction

The first slide is usually the title slide. Remember that the title of your talk need not match the title of the paper or source you are presenting. Also, it is fine not to put title on all slides.

Follow your title slide with a short introduction to the problem. In a longer talk, consider a contents slide *after* the introduction, to give the audience some perspective. Your presentation should tell a story, but it is important that the audience learns the main point of the story *early*, so that they can use it as a frame of reference for the other ideas.

6.5.2 Conclusion and Acknowledgments

You should always include a conclusion slide; leave it up while answering questions. You should also have slides with acknowledgments and references. It can work to place the acknowledgments at the beginning; this gives the audience an opportunity to get used to your accent, without risking that they miss a major point.

6.5.3 Managing Time

Depending on how much time you have, you should decide the approximate number of slides you need. It takes around 1.5–3 minute per slide if you are trying to explain a major point. Title and conclusion slides take much less time. So, a 10 minute presentation has between 5 and 10 slides. However, you can have more slides if you explain a key idea by a series of diagrams or animations where each slide is viewed for a few seconds only.

Draft your slides on quarter pages or index cards; this helps gauge the size of the slide. Not every slide needs a title; consider using title to highlight a change of topic. Plan some flexibility into your talk so that you can skip a few slides to fit the talk into the available time without affecting your rhythm or omitting key ideas.

Depending on your target audience, provide context. Emphasize results and techniques from a variety of sources. Expect to use a non-uniform level of detail pertaining to the interest of the audience. That is, select a one or two ideas to examine in detail, and others to skim over.

6.5.4 Conveying the Key Ideas

Knowing what you want to convey is important, but it is equally important to understand how to convey it so that the audience understands your key ideas. Do not write more than 4-8 lines per slide. Wherever possible, use a graphic (rather than words) to explain the key idea of the slide. On text slides, use visual aids like color, indentation, text highlighting, and varied line lengths to make the content more comprehensible for an audience member taking who is taking a cursory look. Abbreviations and sentence fragments are fine in presentations; they actually help the user to follow and

read quickly.

Involve the audience by asking a rhetorical question, or a question related to common experience. Give explanations and definitions via example. Remember that the audience only has a few seconds to look at your slide so you need to highlight what they are supposed to see. Build in "re-entry points": places where a listener can pick up the thread again. Remind people of a definition if several slides have passed between its introduction and use.

Keep in mind that the purpose of any presentation is to help the audience understand your key idea rather than to impress them with your knowledge of your presentation tool. So, keep your slides simple and clear. Graphics should be straightforward and uncluttered; make sure that there is a "grammar" to the graphical components. For example, all the lines should be of the same width and color, unless there are really two different kinds of line, in which case the two kinds should be distinguished in both width and color.

Try to give the audience a visual outline of the story. Put visually-interesting features, such as diagrams, tables, graphs, or photographs, on at least one-third of your slides. If you are including diagrams from other sources, decide if you need to re-draw them for the presentation, so that they can be of an appropriate size and style. Use fonts large enough to be seen by the audience. The best font size to use depends on the font, as well as the size of the room in which you will be presenting.

If there is a reason to put more than one idea on a single slide, use color to differentiate them. Remember that about one-quarter of North American males are red-green colour-blind, so don't depend on the audience being able to see the difference between small areas of red and black. Yellow and pastel shades can disappear against a white background, so if you use these colors, outline them in black. Similarly, red can disappear against a black background.

Finally, always number your slides. It helps enormously when it comes to question time. It's also very useful when your colleagues are helping you to improve you talk.

6.5.5 Directing Attention

When explaining a concept in a conversation, you might move over to a whiteboard to draw a picture. You don't overwhelm the other person by drawing the whole picture at once. Instead, you draw a fragment of the picture and direct the other person's attention by limiting explanations to this fragment. After they understand the current fragment, you extend the drawing with the next fragment and explain that. Introducing a big concept in pieces keeps everyone on the same page, and makes the completed picture

understandable.

When you are giving a presentation, even though your visual medium is likely to be slides rather than a whiteboard, it pays to think about the conversational whiteboard metaphor when designing figures. If you present a complex figure all at once in a single slide, you will shock the audience with too much information. Instead, break up the figure into fragments and use a series of slides, or animations, to add one fragment at a time. This way the audience can match up what you are currently explaining with the current fragment of the figure. You can even "gray out" everything but the current fragment, and then have a final slide where all the gray has been removed from the figure.

6.6 Practice the talk

Practicing the talk will not only give you an estimate of how long it is, but will also help you to evaluate its coherence and organization. The best way to practice is to deliver your talk out loud in front of a mirror, or by video recording it. Video-recording has the advantage that you don't have to deliver the talk and evaluate it at the same time; it also lets you share your video with colleagues and family for feedback. Using a mirror is simpler, demands less equipment, and is faster when you are fine-tuning the talk.

As a result of practicing your talk, you may decide to:

- reorder slides for better flow;
- add a slide to connect parts of the talk that seem disconnected;
- reduce the number of slides, or the content of some slides, to keep within the time limit; and
- add figures or examples to help explain an idea.

Did you find yourself waving your hands or drawing in the air? If so, add a figure so that you can point to a slide instead! Did you find yourself spending much longer than you expected on a slide, explaining an idea that was not explicitly called out? If so, you should either skip over that idea entirely, or add a prior slide explaining the idea explicitly. You will probably find that adding a slide of explanation will take less time than having to extemporize over it.

Try to anticipate questions. Identify the parts of your presentation that may be hard for the audience to understand, and try to think of the questions that they may ask. Prepare backup slides that help you answer these quesitions.

6.7 Giving the Talk

6.7.1 Before the talk

What better way is there to make a bad first impression than to show up late, not know how to run the equipment, and then have to rush through the rest of your presentation to make up for lost time? Don't let this happen to you!

Build familiarly with the venue. If you traveled to give the talk, find out where the venue is on a map, an know how long it will take you to get there from the hotel. Will that be affected by rush-hour traffic? If you have to park a vehicle, how long will that take.

Visit the site of the talk beforehand — certainly a few minutes early, and, if possible, the day before. Even if you are familiar with the venue, still come early, especially if you are the first presenter or the day. If something is broken, you will have a time buffer to use getting the equipment repaired. How does the audio system work? Is it a microphone and podium arrangement, or will you be equipped with a wireless microphone? How about the video equipment? Operate the equipment and make sure that it works — this is also a good opportunity to practice part of your presentation. Check that the projector has a working bulb. Is there a spare bulb? I've seen it suggested that you make sure that the spare bulb works too, but in practice you will not want to disassemble the projector in someone else's venue. Is the data projector 4:3 ratio or 16:9 ratio? Can you pre-configure your computer so that your slides are not distorted or cropped?

6.7.2 Getting started

Don't start the presentation until you are ready, but start on time. Make sure that the microphone is on, and that you have a clock, or that the time prompter is in sight. Decide where to stand so that you can look at the audience, and not trip over cords when the room goes dark. Prefer sticks, telescopic pointers or the computer's cursor to laser pointers, which tend to be to feeble to be easily visible.

If the talk is long, make sure you have some water accessible. If you think that they may help you, have some throat lozenges in your pocket. If yours is the first talk after a break, make sure that the projector screen is not washed-out by outside or overhead light; if it is, remind the AV person to dim the lights.

6.7.3 Grab your audience

Toastmaster's International says that audience members will decide in the first 30 seconds whether or not they care about what you have to say. You have just that amount of time to connect with the audience, and to set the tone of having a conversation. You can say something intriguing, tell a personal anecdote such as how you first became interested in the topic, tell a relevant joke, ask the audience a question, quote an interesting statistic, or display a relevant prop. Byron Nevis told me that he once saw a talk at OSCON by Ingy döt Net on a data interchange language called YAML. YAML was so flexible that you could use it even if you couldn't decide which language to program in. To make the point, Ingy döt Net came on stage wearing a fisherman's vest with 200 pairs of glasses attached, and switched randomly between glasses throughout the talk. Bryon says that he didn't find YAML very interesting, but that he couldn't pull himself away from the talk for fear of missing the next glasses change!

6.7.4 Use the Available Space

Don't stand rooted in one spot unless you are talking to 500 people in a completely darkened room. As you talk, make eye contact for a solid five seconds with someone in the front row, someone on the side of the room, and someone in the back row. Keep eye contact with different people in the audience throughout your talk.

It's rare to be in a talk where the speaker is too loud. Stand up straight, speak from your diaphragm, and project your voice. Don't be afraid to show excitement in the exciting parts — if you don't show interest in your topic, neither will the audience. If you tend to fidget, hold a pen, a pointer, or a slide-advance clicker, so that you have something to fidget with; it's better than putting your hands in your pockets and fiddling with your keys.

Make use of the space that you have. Use the entire stage, but don't walk around randomly — move around deliberately and with purpose; when you reach a place in the room, stay there for a while before moving somewhere else. If you need to point to something on the far half of a slide, move to that side of the room before that slide comes up.

6.7.5 Directing attention

You have to get the audience to pay attention to the right part of your slides. You can do this in a number of ways. Verbally, you can say "notice that the curve has two bumps." You can also gesture by using your finger, a stick, or the computer cursor. If the screen is low enough, walk up to it and point with your arm or a stick. Or anticipate the need to direct attention, and animate a

label onto the slide.

6.7.6 Using notes

Try not to use notes; your talk will flow more smoothly if you have rehearsed it enough so that you *know* what to say. Notes are just one more thing to distract you from your audience. If you must use notes, hand-written 3×5 index cards are a good mechanism. Alternatively, most presentation software allows you to run dual-screen, with speaker notes facing the speaker, and slides facing the audience. Perhaps you can configure the screen so that the speaker's notes are large enough and prominent enough for this to work for you? Remember, notes are just notes: don't use full text, and don't read your talk. In case you lose the thread when switching slides, consider just one line for each slide with the first thing to say. A little silence is OK; if you need to, stop for a few seconds to gather your thoughts. The audience can use the time to reflect on what you have said so far.

6.7.7 Timing

Have targets for the time in your talk by which you should reach certain slides. For example, use notes on your slides to remind yourself where you should be 50, 75 and 90 per cent of the way through the talk. Use these notes to help you regulate your rate of speaking. If you find that you are drifting from your target, adjust your talk on the fly to get back on schedule, either by omitting material you identified as optional, or adding an example or anecdote. If you tend to speak too fast because of nervousness, use speakers' notes to remind yourself to slow down.

In spite of all of this planning, sometimes you will completely loose track of time; this can easily happen if you include a live demonstration rather than a movie. If you hear the time keeper say "Five minutes!", what should you do? Don't just start talking faster, hoping to cram the remaining 20 minuets of your talk into 5! Do you have a plan for this eventuality? If not, take a moment to stop and think: How can I recover? What should I omit? What's the best use of the time I have left?

An old trick for finishing on time — and you should *always* finish on time — is to know that your conclusion will take x seconds. Then, x seconds before the time limit, jump to your conclusion slide. Never complain about lack of time during your talk — it is a sign to the audience that you are not prepared.

6.7.8 Mistakes

Mistakes happen. The best way to handle a mistake is to publicly acknowledge the it by verbalizing the event, or even making a joke out of it. "Oops! I dropped my notes! But that's better than dropping my computer!" Trying to cover up a mistake will most likely fail; in the process you will insult your audience. Instead, respect your audience, admit your mistake, recover as gracefully as you can, and move on.

6.8 Answering questions

At the end of a presentation, reserve time for the audience to ask questions. Some presenters attempt to avoid the question and answer session since they are worried that they may not know the answers. This is a mistake, because the question and answer session is often the most exciting part of a presentation. It's also the part that most closely approaches a personal conversation. Answering questions gives the speaker a chance to communicate with the audience directly, and further clarify and consolidate the "take aways" from the talk. So treat it as an indispensable part of the presentation — one which deserves as much planning and control as the delivery of the core material.

Prepare for questions. You should prepare yourself for the question and answer session when you prepare the presentation. Naturally, you should understand everything in your slides, including figures from other sources. List possible questions, craft the answers and rehearse them before the presentation. Get ready for difficult questions. Practice answering questions with a confident and a positive attitude. If you anticipate communication difficulties, for example, in a large room without microphones for the audience, consider having some small pieces of paper available for people to write down their questions and pass them to you.

Listen carefully to the *entire* question. Don't interrupt people before they finish the question. Maintain eye contact to show that you are focused on the question, and check that you understand the question before starting to answer it by summarizing the question for the audience, and asking for confirmation from the questioner. It's always a mistake to rush in to respond to the question before it is fully stated; you may think that you are saving time, but if you misunderstand the question and provide an unrelated answer, you are actually wasting time.

Think about the question before responding. Unless there is a simple factual answer, allow yourself some time to evaluate the question and formu-

late a response. A pause is OK; the audience will not expect an immediate answer to a difficult question. At the same time, you can comment on the question, such as "That is a good question" — but don't repeat the same comment after every question!

Answer the question. Try to answer a question precisely and honestly. Answer the question that was asked, not some related question. If you are not sure of your answer, say so, for example, you might say "It is my understanding that . . . '. If you don't know the answer to a question, say so! This is much better than pretending that you know, or giving an unrelated answer: these tactics will fool no one.

Keep your answers brief; avoid going into too much detail or engaging in a long dialog. If the questioner wants to discuss more in detail, you can remind them that other questions are waiting and suggest that you talk after the session. It's better to allow time for another question than to drift off into a long discussion on the previous one.

Talk to the whole audience while answering the questions and make sure everybody can hear your answer — not just the person who asked the question. As you complete your answer, look at the questioner and ask whether your answer is sufficient, or check their body language.

Handling tough questions. If you don't know an answer to a difficult question, treat it with a calm and positive attitude. For example, you might say, "That's an interesting idea; I've never thought about it", which sounds more enthusiastic than saying "I don't know." If you know there are experts in your audience, you can ask if they can answer the question. You can also write down the question, research the answer and follow up with the questioner later.

Avoid an aggressive or defensive attitude when you are asked a hostile question. Politely thank the person for the question, reply with a firm and collected tone, and move on to the next question.

Maintain control. Stay in control of the entire question and answer session. Clearly announce the start of the question session, and explicitly invite questions. Allow only one person to speak at a time. Limit the time you spend answering each question, and solicit the next question. Avoid answering questions that fall outside of the scope of your presentation.

End on a high note. If you can, wrap up the question and answer session with a closing statement. You can summarize your main idea and reinforce your ideas. Don't end your presentation by asking "Any other questions? I guess that's all for today" or by letting the question and answer session

peter out. Instead, close the presentation with a strong ending, and make a graceful exit.

In summary, remember that the question and answer session is just as important as the presentation itself: you should plan it carefully. Involve the whole audience when answering questions. Keep your answers specific and brief, and don't be afraid to admit when you don't know. Stay positive and enthusiastic, and try to end with a conclusive statement.

Reading makes a full man,
meditation a profound man,
discourse a clear man.

Benjamin Franklin,
Poor Richard's Almanack, 1732–57

The art of conversation is the art of hearing
as well as of being heard

William Hazlitt, "On the Conversation of Authors",
The Plain Speaker, 1826

Acknowledgments

Bibliography

[1] I. Alagiannis, R. Borovica, M. Branco, S. Idreos, and A. Ailamaki. Nodb: efficient query execution on raw data files. In *Proceedings of the 2012 international conference on Management of Data*, pages 241–252. ACM, 2012.

[2] Lyn Dupré. *Bugs in Writing: A guide to debugging your prose*. Addison Wesley, 1998.

[3] H.W. Fowler. *A Dictionary of Modern English Usage*. Oxford University Press, Oxford, New York, second edition, 1985. Revised by Sir Ernest Gowers.

[4] D. E. Ingmanson and W. J Wallace. *Oceanography: An Introduction*. Wadsworth, Belmont, CA, 1985.

[5] Oracle Corp. JVM tool interface (version 1.2). http://docs.oracle.com/javase/6/docs/platform/jvmti/jvmti.html, 2011.

[6] J. Rodgers. The life history of a mountain range — Appalachians. In K. J. Hsu, editor, *Mountain Building Processes*, pages 229–243. Academic Press, Orlando, 1983.

[7] William Strunk. *Elements of Style*. Press of W.P. Humphrey, Geneva, N.Y., 1918. Republished by Bartleby.com, 1999.

[8] Mary-Claire van Leunen. *A Handbook for Scholars*. Oxford University Press, 1992.

[9] G. E. Vink, W. J. Morgan, and P. R. Vogt. The Earth's hot spots. *Scientific American*, 252(4):32–39, 1985.