Scholarship Skills

Andrew Black
PSU

Lecture 5: Mechanics of Writing

All material © 1996–2007
David Maier, Todd Leen, Andrew Black
Mechanics

• Grammar, Punctuation, Technical Notation
• Why does it matter? Poor grammar:
  – distracts
  – misleads
  – slows the reader down

• We will criticize some constructs that are grammatically correct, but not optimal (clarity, conciseness etc.).
Spelling

• Use a spell checker
  – But be careful of the corrections

• If you are writing for a US conference or journal, use American (rather than British) spelling
  – “color” vs “colour”
  – Conversely for a European venue?

• Be consistent
  – don’t use “modeling” and “modelling” in the same paper.
    (Core rule: use a consistent lexical set.)

• Avoid contractions in technical writing
  – doesn’t → does not
Word Choice: Confusables

• *affect* vs *effect*
  – *affect* is a verb; *effect* is a noun
  – Doubling the running time had no _____ on the quality of the solution.
  – Doubling the running time did not _____ solution quality.

• *stationary* vs *stationery*
  – *stationary* = traffic doesn’t move;
    stationery = stuff to write on (e is for ✉)

• *complement* vs *compliment*
  – to complete or round out (“e” is round)
Confusables 2

• alternate vs alternative
  – alternate = every other; alternative = a substitute
    The procedure rounds up and rounds down on _____ iterations.
    If the search fails, it tries an _____ strategy.

• imply vs infer
  – authors imply, readers infer
    Douglas and Watson _____ in their paper that method A is superior.
    (See Dupré Segment 64)
Confusables 3

• less (quantity) vs fewer (count)
  – Algorithm A took ____ time than Algorithm B, but B used ___ disk pages.
  – English distinguishes between count nouns, and mass nouns
    How many bagels?
    How much food?   (enough for an army)
    How many foods?  (Milk, cheese and eggs)
    How much money?  (gold? space? water?)
  – Mass nouns are grammatically singular
Latin

• Avoid Latin.
  – e.g. → for example, for instance
  – i.e. → that is
    (but both of these are ok in parenthetical remarks.)
  – etc. → and so forth (but it’s better to explicate examples).

• Black uses Latin if there is no English equivalent.
  – For example, *sic*, [literally “thus, so”] used to mean that the preceding obvious error was indeed part of the the original in a quoted passage.
    e.g., The House of Representatives shall chuse *[sic]* their Speaker and other Officers; and shall have the sole Power of Impeachment.

• The only Latin phrase that Maier uses regularly is “et al.” in citations.
  – However Dupré (Segment 65) suggests that you should not use “et al” at all! Instead, she suggests you use “and colleagues”, or “and associates”.
More on *et al.*

- *et al.* is a contraction of the Latin for “and others”
  - ‘*et alii*’ (masculine plural) or ‘*et aliae*’ (feminine plural) or ‘*et alia*’ (neuter plural)
  - Hence, *et* is not an abbreviation, and does not need a period, whereas *al.* is, and does.

- In a reference list, it is usual in CS to list all of the authors, so there is no need to use *et al.*

- In running text, use “*et al.*” or “and colleagues”: you choose, but be consistent.
  - Dupré and associates [1978] reported severe snoring during the 1977 meeting.
  - Dupré and colleagues [25] reported severe snoring during the 1977 meeting.
  - Dupré *et al.* [1978] reported severe snoring during the 1977 meeting.
  - Dupré *et al.* [25] reported severe snoring during the 1977 meeting.

- If you use author-date citations, and *et al.* in the reference list, then use it in the text too.
  - Consistent lexical set
Jargon

• We often don’t realize that we are using jargon from CS
  – We sought user input on the design.

• Non-computer-scientists can help you spot jargon

• Technical terms used in their correct technical sense are fine.
Numbers

• Dupré segments 24 (Number Spelling) and 34 (Number Styles)
• Generally, spell out whole numbers less than 10
  – We have three choices for each trial.
• Exceptions
  – units of measure, time, dates, page numbers, chapter numbers, percentages, money, proportions, part of a series of larger numbers
    The algorithm failed to converge for $n$ equal to 7, 14 and 21.
    Point A is 3 inches from point B.
  – when two numbers are together, if confusing
    eleven 1000-ohm resistors
  – Time
    He arrived at 8:30pm.
• Use a leading zero on a fraction
  – 0.5 second, but:
  – one-half second, if not accurate to a tenth
Numbers, Cont.

• Use a hyphen for part of an adjective:
  – 5-inch ruler

• Ordinals:
  – 0\(^{th}\), 1\(^{st}\), 2\(^{nd}\), 3\(^{rd}\), 4\(^{th}\)

• Tables: keep a constant number of significant figures in a column.
2 Is Not 3

• Comparatives and Superlatives
  – -er, more for comparing 2 things
  – -est, most for comparing 3 or more things
    Algorithm A is the better of the two.
    Of the ten algorithms we analyzed, Mergesort is the fastest

• Between (2) and Among (3)
  – The procedure places a space ____ each pair of words.
  – The procedure must choose ____ the empty buffer slots.

• “Between” is often used for both in modern writing; some authors consider “among” to be archaic.
  – (For this use of quotation marks, see Dupré Segment 41.)
Citations

• Don’t use citations as nouns
  – In [15], the authors extend the method to lists.

• Instead, treat them as parenthetical remarks
  – McDonnel and Slington [15] extend the method to lists

• Don’t use citations in titles and section headings (or abstracts?)
  – 2.4 Dijkstra’s Algorithm [17] for Shortest Path
In case you missed it:

Do not use citations as nouns
Other Words

• Above and below
  – should appear after the noun
  – ... in the formula above
  – ... in the formula below

• Such as: possibly the same, or as an example
• Like: similar but not equal
  – To order the result, we can use a sorting routine _______ Quicksort.
  – There is no result ___ the pumping lemma for recursive languages.
    Think of: “that is like”, “who are like”

• See Dupré Segment 27: “Like Versus Such as”
Grammar

• Singular and Plural
  – Subject and verb agreement—know what the subject is
    These problems arise …
    This problem arises …

    A subset of these problems arises in the deterministic case.

• Referent of a pronoun should agree in number
  – Footnotes are a means to convey incidental information to the reader. Use them sparingly.

    The footnote is a means to convey incidental information to the reader. Use _____ sparingly.

• Articles: A, An, The
  – Do it by sound. Use “an” before a vowel sound, “a” before a consonant sound.

    An honorable mention, but a unique house
  – See the Useful Links page for a resource on using articles
Avoid Non-referential “This”

- Reducing the number of service queues increases average delay and reduces the number of idle periods. This affects the recovery subsystem.

  • Almost always clearer if you put a noun after “this”

- This reduction affects the recovery subsystem.

  • Use the same term (Consistent Lexical Set)
Keep Parallel Clauses Parallel

– We can solve the problem by making the input file name the default or select the previous output file when none is specified.

• If in doubt, make two sentences in your head:
  – We can solve the problem by making the input file name the default.
  – We can solve the problem by selecting the previous output file ...

• Also use parallel construction in adjacent sentences when they express related (or contrasting) ideas.
Placement of “Only”

• Different placements of “only” change meaning.
  – Negative edge weights affect only Dijkstra’s algorithm.
  – Only negative edge weights affect Dijkstra’s algorithm.
  – Yesterday I hit the man in the eye

• The location of “only” defines the meaning
  – See Dupré Segment 5
Lists of Examples

• Generally, use “and” as the conjunction
  – Examples of computational chemistry codes are GAMESS, Gaussian, and MELDF.

• Don’t use “for example” (or e.g.) together with etc.
  – Mosaic supports multiple image formats including JPEG, PICT, GIF, etc.
  – Mosaic supports multiple image formats including JPEG, PICT and GIF.
“Which” vs “That”
Descriptive vs. Restrictive Clauses

– The lawnmower, which is in the garage, is broken.
– The lawnmower that is in the garage is broken.

– The dinner, which was prepared by an expert chef, began at 9:00pm.
– The dinner that was prepared by an expert chef began at 9:00pm. (This suggests that other dinners, prepared by novices, started at other times.)

• Observing the distinction can remove ambiguity.
• The mistake is almost always “which” for “that”.
• If there is no preceding comma, then you probably want “that”.
  – Which requires a preceding comma because it must introduce a parenthetical remark.
• See Dupré Segment 17.
Split Infinitives

• An infinitive is the “to” form of a verb. For example: “to shoot”, “to drive”

• A split infinitive occurs when you place text between the “to” and the verb:
  – This event is sure to quickly degrade performance of the kernel.
  – The parser must be re-entrant in order to correctly handle macros.

• Dupré argues strongly against splitting infinitives. Some argue that it is OK.
  – If you’re tempted to do it, make sure that a different construction doesn’t flow better.

• See Dupré Segment 38 and Fowler “Out of the frying pan”
Punctuation

• Capitalization
  – Capitalize section, chapter, figure when used as proper names.
    “in Section 3.2” “from Chapter 4”
  – but
    “the previous figure”
    “the next chapter”
  – Optionally capitalize after a colon if and only if what comes after could be a complete sentence. (Correct the following.)

    The result is always the same: reduction in the number of queues increases running time.

    The result is always the same: Increased running time.

    He could see the improvement: The subject now completed the tasks easily.
Capitalization Cont.

• Recast a sentence so it does not begin with a lower case letter (including Greek letters).

  – $a$ and $b$ must be palindromes.
  – Both $a$ and $b$ must be palindromes
  – $ackermann$ is a recursive function.
  – The function $ackermann$ is recursive.
Solidus (Forward Slash)

• According to Dupré, most US publishers do not allow the solidus in text. She suggests that you avoid using it and gives alternative constructions.

• Don’t use the solidus for equally-weighted pairs; use the en-dash instead
  – input/output $\Rightarrow$ input–output, but
  – m/s (for meters per second) is correct

• Don’t use and/or, his/her.
Solidus (Forward Slash)

• Don’t use a slash instead of a conjunction or preposition.
  – The user can then accept/reject the proposed correction.
  – The user can then accept or reject the proposed correction.

  – This bug/feature has existed since Release 2.0.

• and/or is ugly, and ambiguous
  – The exceptional cases can be retried and/or stored to a file.

• See Dupré Segment 117
Parentheses

• See Dupré Segment 37.
  – Enclose stuff in parentheses that the reader can skip the first time through. Parenthetical remarks provide secondary information (stuff that is less important than the main text).

• When words inside parentheses constitute a full sentence, the period goes inside the parentheses.
  – The waveform will be aperiodic (except when the voltage is zero).

  – The waveform will be aperiodic. (The case where the voltage is zero is an exception.)
Quotes

- See Dupré Segment 41.
- By a printers’ convention, punctuation at the end of a sentence moves inside quote marks.
  - Conventional
    The manual states that the meaning of error messages “should be obvious from the context.”
  - Pedantic
    The manual states that the meaning of error messages “should be obvious from the context”.

- Exceptions:
  - One or two words in quotes
    After running the garbage collector, we may still have values that have “spoiled”.
  - The quoted material is for a computer
    After the prompt, type “mv a.out myfile”.
  - Quotes that are used to signify that the word is the object
    It’s best not to use contractions like “can’t”.
Colon

• See Dupré Segment 15
• “A colon signifies that what follows it expands on or explains what precedes it: this sentence is an example.”
  – (Note that the construct following the colon can be a sentence fragment rather than a complete sentence.)

• You might use a colon instead of a period when the following sentence explains the first. An em-dash (the widest dash “—”, obtained in TeX by “---”) also works.
  – Using a period is grammatically correct, but the colon (or the dash) makes the connection between the sentences explicit.
Colon

• Example from Dupré

  Lyn could tell that BB was going into heat. BB had been howling and yowling the entire night through.

  or

  Lyn could tell that BB was going into heat: BB had been howling and yowling the entire night through.
Semicolon

• See Dupré Segment 93
• “The semicolon connects two sentences that are closely related to each other.”
• What follows must be a complete sentence, not a sentence fragment (string of words that don’t constitute a sentence).
  – When what follows is a sentence fragment, use a comma, or an em-dash, or a colon.
Semicolon

• Examples from Dupré:

  – This machine is difficult to use; it crashes whenever you change windows.

  – Holly and Misha cooked yet another humongous meal — and refused to let anyone help clean up.

  – Max’s heart was throbbing; Lyn’s head was swimming.
Hyphen

• A hyphen is used to glue two words together to form a compound *adjective*.

• Adjectives (including nouns as adjectives) are right associative unless overridden by a hyphen.
  – large object repository (large repository for objects)
  – large-object repository (repository for large objects)

• Except if the modifier is a commonly known phrase
  – operating systems tutorial

• Words that describe *activities* don’t need hyphens
  – Decision making under uncertainly is difficult
    *decision making* is an activity
  – Decision-making algorithms are hard to formalize
    *decision-making* is a compound adjective

• See Dupré Segment 29
EM-dashes —

• See Dupré Segment 77
  – “An em dash — the punctuation-mark dash — delineates a thought that is not a critical part of your sentence.” (It functions like parentheses.)
  – “… can function as a strong comma, or as a semicolon; it is, in essence, a loosely defined break or pause in a sentence that calls attention to the material that it sets off.”

• Learn how to use your tools!

<table>
<thead>
<tr>
<th>Symbol</th>
<th>name</th>
<th>Typewriter &amp; LaTeX</th>
<th>Macintosh</th>
<th>Windows</th>
<th>Linux</th>
<th>HTML</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>hyphen</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>−</td>
<td>en-dash</td>
<td>--</td>
<td>\ -</td>
<td>Alt+0150</td>
<td>Compose+ --</td>
<td>–</td>
</tr>
</tbody>
</table>
| —      | em-dash  | ---               | \[
     |          |        |           | Alt+0151 | Compose+ --- | &mdash; |

Windows character codes must be typed on the numeric keypad with Num Lock on.
em-dashes —

• Examples from Dupré:

  Minsky, Simon, Newell, and McCarthy — all grand men — decided to play a genuinely intelligent game.

  or

  Minsky, Simon, Newell, and McCarthy (all grand men) decided to play a genuinely intelligent game.

  The em-dashes emphasize the parenthetical remark, the brackets de-emphasize it.

  Shannon — the founder of information theory — gave an interesting statistical account of information content in written English.
EN-Dashes –

• Dupré Segment 49
• Used to indicate ranges; pronounced as “to”
  – Read pages 5–23
• But it is better to spell things out
  – Read pages 5 to 23.
• in text, and use the en-dash to indicate ranges in tables (and their titles), figures (and their captions), and page ranges given in references.
en-dashes –

• Also use the en-dash to join an equal-weighted pair used as an adjective

  – Have you attended a life—work planning seminar?
  – This requires a device specific driver.
  – This is the on—off switch
  – Theirs was a on-again—off-again relationship.
  – The New York–London flight arrived late

  Does this mean that the flight from New York to London arrived late, or that the new flight from York to London arrived late?
if ... then

• Always put a comma before then.
  – If the speaker pauses, then the word might not be recognized.

• Even if there is no then: dependent clauses end with a comma.
  – If we close the application at this point, the changes will be lost
  – Some authorities warn not to omit the “then” after the dependent clause.
Past vs. Present Tense

• Use the same tense throughout a discussion.

• Past Tense
  – Use past tense to indicate that results apply only to the particular study or experiment.
    The survey showed that this population believes …
  – Use past tense to say what you did (for example, in an experiment).
    We removed all personal information from the data
  – Use past tense to point to a result earlier in your paper
    In Section 2 we showed that …
Past vs. Present Tense

• Present Tense
  – Use present tense to indicate that results generalize:
    The experiments show that loops degrade the performance of Matlab programs significantly.
  – Use present (or future) tense to say what you are going to discuss later in a paper.
    In the next section, we show … or
    In Section 4 we will show
  – Hypotheses, theorems, and facts are expressed in present tense.

• These rules also work if you’re discussing the work of other authors.
Subject-Verb

• Put no comma between subject and verb
  – All of the students in the class arrived in time

• Unless it is there to set off a clause:
  – All of the students in the class, along with many who were not registered, arrived to eat pizza

• This rule means that there will be either 0 or 2 commas between subject and verb!
Fonts

- Fonts can help or confuse the reader
- Font is important when a name or term is an ordinary word
  - The next token will be an, and or or, but not not.
- Decide what fonts mean and be consistent
  - variables
  - code
  - algorithms, systems, operators
  - When factor is on iteration i of the main loop, the statement $x := f(x)$ will have executed at least $i - 2$ times

• Typewriter font is ugly and hard to read: there are better choices.

• Create a style sheet
  - Define styles or macros in your document tool
Fonts

• Fonts can help or confuse the reader
• Font is important when a name or term is a ordinary word
  – The next token will be an, and or or, but not not.
• Decide what fonts mean and be consistent
  – variables
  – code
  – algorithms, systems, operators
  – When factor is on iteration i of the main loop, the statement
    \( x := f(x) \) will have executed at least \( i - 2 \) times

• Typewriter font is ugly and hard to read: there are better choices.
• Create a style sheet
  – Define styles or macros in your document tool