Lecture 8
Typesetting

Prof. Dr. Jasmin Blanchette

Chair of Theoretical Computer Science and
Theorem Proving

Version of October 21, 2025

General Principles

Font Families and Sizes

Named collections of related fonts (for bold, italics, etc.) form a font family—e.g., Times New Roman.

Use **as few font families** as possible while making important distinctions visible.
Often a single family suffices for an entire document, including diagrams.
You can make an exception for code, which is usually set in a monospaced font.

Combine font families only if they have similar line thickness and
the lowercase letter x has a similar height in all of them.

Also ensure that your text is set in a suitably **large font**.
This principle applies especially for slides and posters.

Bold and Underline

Emphasis can be achieved using bold, underline, italics, and colors.

Bold is often used for **titles** and **headings**.
It can be used for emphasis on slides and posters.

Keywords in pseudocode are often set in bold—e.g.:

if $x < y$ **then**
  $x := y$
else
  $y := x$

Underline—e.g., *Journal of Cryptology*—is **ugly**.
It was used on typewriters when italics were not available. Prefer italics if possible.

Italics

Italics are elegant, but they **slow down** reading.
Beyond emphasis, they have three main uses.

- Italics highlight **uncommon foreign phases**, as in *This thesis is a veritable* tour de force. Do not italicize common Latin phrases such as *e.g.*, *et al.*, *etc.*, and *i.e.* In case of doubt, check your dictionary.
- Italics can be used to **define a term**, as in *A list whose reverse is equal to itself is called a* palindrome. Italics should not be used for other purposes nearby.
- Italics are used, in conjunction with headline-style capitalization, for the **titles of "large works"** such as theses, books, journals, and music albums when they occur in text—e.g., *This paper was submitted to the* Journal of Cryptology.

Within an italicized passage, you can emphasize a passage by unitalicizing it.

Colors

Colors can be used for emphasis or to make distinctions.
In programming, they are frequently used for **syntax highlighting**.

Bear in mind that some people are **color-blind**.
They often cannot distinguish between red and green, among others.
Thus, avoid relying on color alone.

If you use colors, check that your document is readable when printed in black and white. In addition, choose your colors with taste.
Choose colors that offer a good contrast with the background and with each other.
Avoid pure primary colors, and consider using **palettes recommended by experts**.

Headings

Chapter, section, and subsection headings should be viewed as **meta-information**. The text following a heading should make sense if the reader skips the heading. Compare:

2.2 First-Order Relational Logic

This logic, implemented in Kodkod, is used . . .

2.2 First-Order Relational Logic

Kodkod's underlying formalism, first-order relational logic, is used . . .

Capitalization

Titles and headings can be capitalized following two approaches (van Leunen).

- **Headline-style capitalization**: "Capitalize the first word, the first word after a colon, and all other words except articles, unstressed conjunctions, and unstressed prepositions"—e.g., *An Intermezzo: Jazzing Up the Data*.
- **Sentence-style capitalization**: "Capitalize the first word, the first word after a colon, and no other words except proper nouns and proper adjectives"—e.g., *An intermezzo: Jazzing up the data*.

For headline-style capitalization, consider words with an intraword hyphen as one word, and words with an interword hyphen as two separate words—e.g., *A Post-processor for the Lambda-Calculus*.

For the title of your document, use headline-style capitalization.
For chapter, section, and subsection headings, use either style, but be consistent.

Capitalization

Should you capitalize phrases such as *the Event Dispatcher Thread* and *Fermat's Little Theorem*? And what about *the Introduction*?

In case of doubt, **prefer lowercase**. Thus, write *the event dispatcher thread*, *Fermat's little theorem*, and *the introduction*. Exceptions can be made for landmark results such as *Fermat's Last Theorem*.

Lowercase is arguably more pleasant to read and is easier to use consistently.

When introducing an acronym or initialism, there is no need to capitalize the unabbreviated expression unless it is a proper noun. Compare:

> *The framework fully supports Multiple-Document Interface (MDI).*
> *The framework fully supports multiple-document interface (MDI).*

Numbers

Generally spell out numbers from **zero to ten** (or twelve),
unless this would look locally inconsistent—e.g.:

> *Below, we compare the performance of all eight configurations on our server.*
> *We used a third-party library for programming the user interface.*

Since sentences should not start with a number (or symbol),
also spell out numbers at the beginning of sentences, or reword to avoid the issue.

Numbers with at least five (or four) digits should use the **thousands' separator**,
either a traditional comma or, following international standards, a small space—e.g.,
1,000,000 or *1 000 000*.

For years and page numbers, the thousands' separator is avoided—e.g., *1999*.

Numbers

In English, the **decimal separator** is the period (.).
Thus, two and a half is written *2.5*, not *2,5*.

There should be a space before units—e.g., *30 s*.
In compound expressions, there should be no hyphen—e.g., *a 30 s timeout*.

The percentage sign, the degree sign, and currency symbols have their own conventions—e.g., *3%*, *90°*, *21 °C*, *$55*, *€50*.

When presenting measurements, you rarely need to show more than three **significant digits**—e.g., *89.7%*. Note that *90.0%* is more precise than *90%*.

There is an important distinction between percentages and **percentage points**. Percentage points denote the difference between percentages. A value of 50% falling by 10 percentage points becomes 40%. By contrast, a 10% fall would yield 45%.

Dates

In American English, dates are written in **month–day–year** order, with commas around the year—e.g., *January 1, 1990,*.

If the day or the year is omitted, no commas are needed—e.g., *January 1990*.

In British English, dates are written in **day–month–year** order—e.g., *1 January 1990*.

Further examples:

> *We met on Sunday, February 14, 2021.*
> *We met again on Sunday, 21 February 2021.*
> *The semester lasts from October 2026 to March 2027.*

Do not abbreviate years. Thus, write *1999*, not *'99* or *99*.

Coinciding Punctuation

When multiple punctuation signs would occur in sequence, **special rules** apply:

- Logically, an abbreviation such as *Co.* should become *Co..* at the end of a sentence, but in practice one period swallows the other.
- American typesetting conventions dictate comma–quote (*,"*) and period–quote (*."*), also against logic. In British English, punctuation that does not logically belong inside the quote is put outside—e.g., *I said 'OK'.*
- In *"'Who are you?' she asked,"* a small space is inserted between the opening double and single quotes for readability, and the question mark swallows the separating comma that would otherwise have followed it.

The conventions governing coinciding punctuation are **outlandish**.
For American English, they are spelled out in van Leunen, pages 60–73.

Cross-References

Cross-references are useful in all sorts of documents, but especially in dictionaries and reference manuals, which are usually read out of order.

Forward cross-references are pointers to a location later in the document—e.g.:

> *We introduce a new functional programming language, λ_{xyz} (Section 3).*

Backward cross-references are pointers to a location earlier in the document—e.g.:

> *When we presented the language's grammar in Section 3.2.1, we silently omitted the ternary operator.*

Notice the tenses: present (or sometimes future in lengthy documents) for forward cross-references, past for backward cross-references.

Incidentally, you will rarely need to tell the reader to **see** a given section or page.

Cross-References

If your text contains lots of cross-references, this might indicate a serious **flaw in your document's structure**. For example, if you have many forward cross-references in Sections 3 to 6 to Section 7, maybe Section 7 should have been Section 3.

"Organization is the most elusive of all qualities in writing. We recognize badly organized material when other people produce it, but spotting our own failures is more difficult. Heavy cross-reference is one of the few hints you'll ever have to guide you; heed its warning" (van Leunen).

Footnotes

Footnotes are suitable for **minor comments** that would disrupt the text's flow and that are nevertheless too important to omit.
They are useful for digressions or discussions targeted only at some of the audience.

In English, the footnote call (e.g., [1] or [*]) generally comes **immediately after any punctuation sign**, even against logic. It is better when the call comes after the period that ends the sentence, where it does not interrupt the sentence's flow—e.g.:

> *Zipperposition's success inspired me to develop Jeha, a new theorem prover.*[2] *Unlike its predecessor, Jeha is implemented in Standard ML.*
>
> [2] *The name Jeha—also spelled Joha—is appropriated from a fictional character.*

Most readers will look at the footnote to see what it is. This will cause them to **lose context**. So there needs to be a good reason for the footnote to exist.

Item Lists

Item lists can be used for listing **parallel items** or **consecutive steps**.
As a positive side effect, they visually break up the text.
They work best with lists of up to about five items and should be used sparingly.

Items start with **bullets** (•) or **numbers** (1, 2, 3, etc.).
Keep numbered lists for situations in which the items' order matters or
you need to refer to individual items later.

Make sure that your items are **symmetric**—e.g., they are all sentences, or *-ing* verbs.

Punctuation must be consistent. Some authors punctuate each item, even if it is not a complete sentence. Others omit the punctuation. Compare:

- *Alnitak;*
- *Alnilam; and*
- *Mintaka.*

- *Alnitak;*
- *Alnilam;*
- *Mintaka.*

- *Alnitak*
- *Alnilam*
- *Mintaka*

Spacing

Two **fundamental rules** of spacing:

1. Put less space between closely related elements,
   and inversely more space between loosely related elements.
2. Try to avoid hardcoding dimensions.
   And if you need to, use multiples of logical units (e.g., baselines, ems).

The first principle applies especially to **mathematical operators**.
Operators that bind more tightly, such as '$\cdot$', need less space around them
than operators that bind more loosely, such as '$=$'.
Thus, $2 \cdot 2 = 4$, not $2 \cdot 2 = 4$.

Paragraph Breaks

The **beginning of paragraphs** should be clearly marked,
through either indentation or vertical space (but not both simultaneously).

Exceptionally, the first line of the first paragraph of a document, chapter, section, or subsection is usually not indented in English.

Examples of Paragraph Breaks

Correct:

> *The toilet paper dispensers in a certain building are designed to hold two rolls of tissues, and a person can use either roll.*
>
> *There are two kinds of people who use the rest rooms in the building: . . .*

Also correct:

> *The toilet paper dispensers in a certain building are designed to hold two rolls of tissues, and a person can use either roll.*
>
> *There are two kinds of people who use the rest rooms in the building: . . .*

Incorrect:

> *The toilet paper dispensers in a certain building are designed to hold two rolls of tissues, and a person can use either roll.*
> *There are two kinds of people who use the rest rooms in the building: . . .*

Line Breaks

Line breaking and hyphenation are largely automated by modern text processors, but they sometimes require **special attention**.

Titles and **headings** should be carefully broken. Compare:

> *Simulation of* n*-Tape Turing Machines Using 2*n*-Stack Pushdown Automata*
>
> *Simulation of* n*-Tape Turing Machines*
> *Using 2*n*-Stack Pushdown Automata*

Avoid breaking references to sections, examples, theorems, etc. The break in *Section 3* is distracting. Avoid it by using a **nonbreaking space** (also called "tie").

Avoid breaking immediately before the **last item** of a short inline enumeration, as in

> *The results of our empirical evaluation are summarized in Figures 3, 4, and 5.*

Page Breaks

Page breaking is largely automated by modern text processors,
but it sometimes requires **special attention**.

In particular, try to avoid these:

- A **widow** arises when a page break occurs right before the last line of a paragraph.
- An **orphan** arises when a page break occurs right after the first line of a paragraph.

Justification

Most prose documents are justified, meaning that the whitespace between words **stretched** or **shrunk** so that the lines reach the right margin.

Ragged-right, whereby no stretching or shrinking takes place, can look better with short lines (e.g., in two-column format).

Tables and Figures

Tables and figures are welcome distractions in most documents:

- Tables are **grids** of textual or numeric data.
- Figures are **diagrams** or **images** illustrating some concepts or visually presenting some data.

Tables and figures can be used

- as **displays**—integrated in the text;
- as **floats**—labeled with a (sentence-capitalized) caption.

Example of a Displayed Table

Another measure of MaSh and MeSh's power is the total number of problems solved. The following table gives the success rate in percent for various combinations of selectors and systems:

	CVC4	*Epar*	*Vampire*	*Z3*	*Any system*
MePo	*38.2*	*40.8*	*41.3*	*40.5*	*48.5*
MaSh-kNN	*47.0*	*50.4*	*51.4*	*48.0*	*60.0*
MaSh-NB	*47.9*	*51.0*	*52.0*	*49.1*	*60.3*
MeSh-kNN	*46.7*	*48.9*	*50.8*	*50.2*	*59.6*
MeSh-NB	*46.8*	*49.0*	*51.0*	*51.3*	*60.2*

We see that the success rate goes up substantially, from 48.5% with MePo to about 60% with the new selectors.

Example of a Floating Table

Table 7: Success rates in percent per selector and system

	CVC4	Epar	Vampire	Z3	Any system
MePo	38.2	40.8	41.3	40.5	48.5
MaSh-kNN	47.0	50.4	51.4	48.0	60.0
MaSh-NB	47.9	51.0	52.0	49.1	60.3
MeSh-kNN	46.7	48.9	50.8	50.2	59.6
MeSh-NB	46.8	49.0	51.0	51.3	60.2

Another measure of MaSh and MeSh's power is the total number of problems solved. Table 7 explores various combinations of selectors and systems. We see that the success rate goes up substantially, from 48.5% with MePo to about 60% with the new selectors.

Tables and Figures

Displayed tables and figures should be **connected** to the text.

Floats should be **referred to** in the text, so that they do not get overlooked.
Floats are especially useful to avoid **page break** issues.

Stylish tables contain **no double lines** and, usually, **no vertical lines**.

Your text should **clarify the semantics** of figures with boxes and arrows.
What is the meaning of a dotted arrow? Of a rectangular box?

Mathematics

Mathematical Prose

Fundamentally, mathematical prose is prose. There may be lots of notations, displayed formulas, and other layouts that introduce additional complexity, but the **basic rules** of style, grammar, and punctuation apply.

Mathematical prose should be **optimized for the reader**.
Carefully chosen notations can help capture key concepts and increase readability.
Let the notations work for you.

But be careful: Compact notations that are useful for readers who are immersed in the topic may be overwhelming for readers who are new to the ideas.

For example, the reverse of a word is alternatively written $\overline{w}$, $rev(w)$, or $w_n, \ldots, w_1$.
Which is best?

Mathematical Symbols

Three fine points:

- Denote **multiplication** of x and y by xy or $x \cdot y$ or even $c \times y$, but not $x * y$.
- Denote **minus** by a minus sign ($-$), which is not the same as a hyphen (-), a short dash (–), or a long dash (—).
- When using **ellipses** in the middle of formulas, repeat the operators on both sides, and align the ellipsis with the operators—e.g., $x_1 + \cdots + x_n$.

Logical symbols belong in logical formulas and should generally be avoided elsewhere.
Thus, write *for all*, *for each*, or *for every* instead of $\forall$,
for some or *there exists* instead of $\exists$,
not instead of $\neg$, *and* instead of $\wedge$, *or* instead of $\vee$, etc.
Avoid *for any*, since it can mean both *for all* and *there exists*.

When you use English for complex logical formulas,
make sure to suggest the **intended precedence** of *and*'s and *or*'s.

Mathematical Variables

Metadiscourse is useful to indicate the scope of variables—e.g.:

> *For the rest of this section, let x be a positive real number.*

Try to avoid multiple **layers of subscripts**.
Given a sequence $t_1, \ldots, t_m$, a subsequence would usually be written as $t_{i_1}, \ldots, t_{i_n}$.
Two strategies to avoid the double subscripts:

- Avoid introducing the original sequence as $t_1, \ldots, t_m$ in the first place.
  Then you can refer to the subsequence as $t_1, \ldots, t_m$.
- Refer to the subsequence as $u_1, \ldots, u_n$.

Avoid introducing variables you **use only once**. This is wrong:

> *A node* n *is internal if it has one or more children.*

Mathematical Formulas

Important or large formulas are usually **displayed**. The remaining formulas are **inlined**. But avoid displaying false formulas if possible.

Traditionally, displayed formulas are **punctuated**—e.g.:

> *This indeterminate is instead handled by the formula*
>
> $$\sum_{n_1=0}^{n_2} S_1 g_{\vec{n}} = \sum_{n_1=0}^{n_2} g_{\vec{n}}.$$

Many computer scientists would omit the period, for consistency with displayed code.

Your text should read smoothly even if you read all formulas aloud as "**blah**," which is essentially what many readers will do, especially on a first reading. Introduce redundancy if it helps—e.g., *the equation* $n = m^2$ instead of $n = m^2$.

Mathematical Environments

Environments such as "Definition," "Lemma," "Theorem," "Example," and "Remark" **emphasize** important material and ease **cross-referencing**.
Traditionally, the statement of lemmas and theorems is italicized—e.g.:

> **Theorem 3.1.** If n is an even natural number, then n/2 is a natural number.

State definitions, lemmas, and theorems before you invoke them.
In this way, **circular arguments** are ruled out by construction.

If possible, number all environments using one **shared counter**—e.g.,
Theorem 3.1 is followed by *Example 3.2*, which is followed by *Definition 3.3*, etc.
This helps the reader follow cross-references.

Lemmas and theorems need **proofs**. These start with the label *Proof.* usually typeset in italics, and end with the box symbol (□) or *QED*.

Computer Science

Algorithms

Algorithms can be used to present the main idea of your code abstractly.
Some advice:

- Explain the algorithm **informally** first.
- State its **inputs**, **outputs**, and **specification** clearly.
- Present it in **pseudocode** in a modular fashion.
  The pseudocode can be displayed or put in a float.
  Use the imperative mood.
- Explain how it works using an **example**.
- State its **properties** (e.g., partial correctness, termination, asymptotic complexity) as rigorously as possible.

Example of an Algorithm

Input: A grammar $G = (V, \Sigma, P, S)$ in Chomsky normal form
Output: A set $W \subseteq V$ of all variables that do not generate the empty language

$W := \{A \in V \mid A \to a \in P,\ a \in \Sigma\}$;
repeat
  $W_{old} := W$;
  $W := W_{old} \cup \{A \mid A \to BC \in P,\ B \in W_{old},\ C \in W_{old}\}$;
until $W = W_{old}$;
return W;

Code Excerpts

Small code excepts, of up to one line, can be **inlined** in your text.
Larger code excepts can be **displayed** or put in a **float**.
Make sure to use a special (usually monospaced) font to make the code stand out, whether it is inlined, displayed, or floating.

Unless you are writing software documentation, do not show much code, if any.
You will probably struggle to explain the code, and chances are that the reader will not understand it anyway. It is usually better to **keep things abstract**.

Software

Typesetting Software

Word processors such as Microsoft Word, LibreOffice Writer, and Google Docs can be used to write theses and papers, and **presentation software** such as Microsoft PowerPoint and Apple Keynote can be used to prepare slides and posters.

LaTeX is an alternative to all of the above that is preferred by many computer scientists, especially in academia. LaTeX documents are text files containing typesetting commands. The files are translated to PDF.

Two variants of LaTeX are particularly useful:

- **Overleaf** is an online interface for LaTeX. It is convenient and simplifies using LaTeX because you do not need to install any packages.
- **LyX** is a "what you see is what you mean" graphical interface on top of LaTeX. It encourages you to focus on the structure of your documents, as opposed to their appearance.

LaTeX Example

```
\documentclass{letter}

\begin{document}
To the editors of the \emph{Bulletin of the AMS},

We are writing to report that a direct search on the CDC~6600 yielded
\[27^5 + 84^5 + 110^5 + 133^5 = 144^5\]
as the \textbf{smallest instance} in which four fifth powers sum to a fifth
power. This is a counterexample to a conjecture by Euler that at least $n$
$n$th powers are required to sum to an $n$th power, $n > 2$.

Yours truly,

L. J. Lander and T. R. Parkin
\end{document}
```

LaTeX Example

To the editors of the *Bulletin of the AMS*,

We are writing to report that a direct search on the CDC 6600 yielded

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5$$

as the **smallest instance** in which four fifth powers sum to a fifth power. This is a counterexample to a conjecture by Euler that at least n nth powers are required to sum to an nth power, $n > 2$.

Yours truly,

L. J. Lander and T. R. Parkin

Selected Literature

- Matthew Butterick, **Butterick's Practical Typography**, Second Edition, `https://practicaltypography.com/`, accessed July 21, 2025.
  This book provides an accessible introduction to typography.
  The author also designs his own fonts, which you can buy.
- Robert Bringhurst, **The Elements of Typographic Style**, Hartley & Marks Publishers, 1992.
  This is another excellent book, which goes into greater detail.