

## SORBONNE UNIVERSITÉ UNIVERSITÉ

ECOLE DOCTORALE INFORMATIQUE, TÉLÉCOMMUNICATIONS ET ÉLECTRONIQUE - ED130

INRIA DE PARIS / ÉQUIPE ALMANACH

THÈSE DE DOCTORAT

Discipline : Informatique

Présentée par

**Jean-Pierre DUPONT**

Dirigée par

**Marie CURIE et Louis PASTEUR**

Pour obtenir le grade universitaire de

**DOCTEUR de SORBONNE UNIVERSITÉ**

---

### **On Modern NLP for Extraterrestrial Languages and Scripts**

---

Présentée et soutenue publiquement le 30 avril 2022 devant le jury composé de :

Alonzo CHURCH	Princeton University	Examineur
Christopher COLUMBUS	Kingdom of Castile	Invited member
Margaret HAMILTON	University of Michigan	Rapporteur
Emmy NOETHER	Georg-August-Universität Göttingen	Examineur
Laurent ROMARY	Inria - ALMANACH	Directeur
Benoît SAGOT	Inria - ALMANACH	Co-directeur
Claude SHANNON	MIT	Examineur
Alan TURING	Princeton University	Rapporteur



## ABSTRACT

Scientific documents often use  $\text{\LaTeX}$  for typesetting. While numerous packages and templates exist, it makes sense to create a new one. Just because.



# CONTENTS

I	A GOOD PART	1
1	INTRODUCTION	3
1.1	Why?	3
1.2	How?	3
1.3	Making this template <i>yours</i>	3
1.4	Features	5
1.4.1	Typesetting mathematics	5
1.4.2	Typesetting text	5
1.5	Changing things	6



# PART I

## A GOOD PART

You can also use parts in order to partition your great work into larger ‘chunks’. This involves some manual adjustments in terms of the layout, though.





# 1 INTRODUCTION

In which the reasons for creating this package are laid bare for the whole world to see and we encounter some usage guidelines.

This package contains a minimal, modern template for writing your thesis. While originally meant to be used for a Ph. D. thesis, you can equally well use it for your honour thesis, bachelor thesis, and so on—some adjustments may be necessary, though.

## 1.1 WHY?

I was not satisfied with the available templates for L<sup>A</sup>T<sub>E</sub>X and wanted to heed the style advice given by people such as Robert Bringhurst ([Bringhurst, 2012](#)) or Edward R. Tufte ([Tufte, 1990, 2001](#)). While there *are* some packages out there that attempt to emulate these styles, I found them to be either too bloated, too playful, or too constraining. This template attempts to produce a beautiful look without having to resort to any sort of hacks. I hope you like it.

## 1.2 How?

The package tries to be easy to use. If you are satisfied with the default settings, just add

```
\documentclass{mimosis}
```

at the beginning of your document. This is sufficient to use the class. It is possible to build your document using either L<sup>A</sup>T<sub>E</sub>X, X<sub>Y</sub>L<sup>A</sup>T<sub>E</sub>X, or LuaL<sup>A</sup>T<sub>E</sub>X. I personally prefer one of the latter two because they make it easier to select proper fonts.

## 1.3 MAKING THIS TEMPLATE *YOURS*

Prior to using this template, the first thing you want to do is probably a little bit of customisation. You can achieve quick changes in look and feel by picking your own

Package	Purpose
<code>amsmath</code>	Basic mathematical typography
<code>amsthm</code>	Basic mathematical environments for proofs etc.
<code>booktabs</code>	Typographically light rules for tables
<code>bookmarks</code>	Bookmarks in the resulting PDF
<code>dsfont</code>	Double-stroke font for mathematical concepts
<code>graphicx</code>	Graphics
<code>hyperref</code>	Hyperlinks
<code>multirow</code>	Permits table content to span multiple rows or columns
<code>paralist</code>	Paragraph (‘in-line’) lists and compact enumerations
<code>scrlayer-scrpage</code>	Page headings
<code>setspace</code>	Line spacing
<code>siunitx</code>	Proper typesetting of units
<code>subcaption</code>	Proper sub-captions for figures

Table 1.1: A list of the most relevant packages required (and automatically imported) by this template.

fonts. With the `fontspec` package loaded and  $\text{\XeLaTeX}$  or  $\text{\LuaLaTeX}$  as your compiler, this is pretty simple:

```
\setmainfont{Your main font}
\setsansfont{Your sans-serif font}
\setmonofont{Your monospaced font}
```

Make sure to select nice combinations of that are pleasing to *your* eyes—this is your document and it should reflect your own style. Make sure to specify font names as they are provided by your system. For instance, you might want to use the following combination:

```
\setmainfont{Baskerville}
\setsansfont[Scale=MatchLowercase]{IBM Plex Sans}
\setmonofont[Scale=MatchLowercase]{IBM Plex Mono}
```

If these fonts exist on your system, your normal text will look a little bit different from the other font used in this example PDF, while your sans-serif font will pair nicely with your monospaced font. You can also remove the `Scale` directive, but I find that most fonts pair better if they are adjusted in size a little bit. Experiment with it until you find a combination that you enjoy.

## 1.4 FEATURES

The template automatically imports numerous convenience packages that aid in your typesetting process. [Table 1.1](#) lists the most important ones. Let's briefly discuss some examples below. Please refer to the source code for more demonstrations.

### 1.4.1 TYPESETTING MATHEMATICS

This template uses `amsmath` and `amssymb`, which are the de-facto standard for typesetting mathematics. Use numbered equations using the `equation` environment. If you want to show multiple equations and align them, use the `align` environment:

$$V := \{1, 2, \dots\} \tag{1.1}$$

$$E := \{(u, v) \mid \text{dist}(p_u, p_v) \leq \epsilon\} \tag{1.2}$$

Define new mathematical operators using `\DeclareMathOperator`. Some operators are already pre-defined by the template, such as the distance between two objects. Please see the template for some examples. Moreover, this template contains a correct differential operator. Use `\diff` to typeset the differential of integrals:

$$f(u) := \int_{v \in \mathbb{D}} \text{dist}(u, v) \, \text{d}v \tag{1.3}$$

You can see that, as a courtesy towards most mathematicians, this template gives you the possibility to refer to the real numbers  $\mathbb{R}$  and the domain  $\mathbb{D}$  of some function. Take a look at the source for more examples. By the way, the template comes with spacing fixes for the automated placement of brackets.

### 1.4.2 TYPESETTING TEXT

Along with the standard environments, this template offers `paralist` for lists within paragraphs. Here's a quick example: The American constitution speaks, among others, of (i) life (ii) liberty (iii) the pursuit of happiness. These should be added in equal measure to your own conduct. To typeset units correctly, use the `siunitx` package. For example, you might want to restrict your daily intake of liberty to 750 mg.

Likewise, as a small pet peeve of mine, I offer specific operators for *ordinals*. Use `\th` to typeset things like July 4<sup>th</sup> correctly. Or, if you are referring to the 2<sup>nd</sup> edition of a book, please use `\nd`. Likewise, if you came in 3<sup>rd</sup> in a marathon, use `\rd`. This is my 1<sup>st</sup> rule.

## 1.5 CHANGING THINGS

Since this class heavily relies on the `scrbook` class, you can use *their* styling commands in order to change the look of things. For example, if you want to change the text in sections to **bold** you can just use

```
\setkomafont{sectioning}{\normalfont\bfseries}
```

at the end of the document preamble—you don't have to modify the class file for this. Please consult the source code for more information.



## *1 Introduction*

## BIBLIOGRAPHY

Robert Bringhurst. 2012. *The Elements of Typographic Style*, 4 edition. Hartley & Marks Publishers, Vancouver, British Columbia, Canada.

Edward R. Tufte. 1990. *Envisioning information*. Graphics Press, Cheshire, CT, USA.

Edward R. Tufte. 2001. *The visual display of quantitative information*, 2 edition. Graphics Press, Cheshire, CT, USA.