

Chapter 2 Typesetting Text

Donald E. Knuth(高德纳)

2017 年 7 月 21 日

1 文章和语言的结构

书写一篇文章最重要的一点是把想法、信息、知识传达给读者。 \LaTeX 与其他类型的排版系统不同,你只需要告诉它一个文本的逻辑结构和语义结构。然后根据文档类文件和各种样式的文件中所带有的“规则”来生成文件。

\LaTeX 中最重要的文本单位是段。分段需要在对应的源码中空一行,如果要继续写可以用换行符,用`\`或`\newline`。下面是说明该换行还是该另起一段的三个正确示例:

`% Example 1`

`\ldots when Einstein introduced his formula`

`\begin{equation}`

`e = m \cdot c^2 \ ;`

`\end{equation}`

`which is at the same time the most widely known`

`and the least well understood physical formula.`

`% Example 2`

`\ldots from which follows Kirchhoff' s current law:`

`\begin{equation}`

`\sum_{k=1}^n I_k=0 \ ;`

`\end{equation}`

`Kirchhoff' s voltage law can be derives \ldots`

`% Example 3`

`\ldots which has several advantages.`

`\begin{equation}`

`I_D = I_F - I_R`

`\end{equation}`

`is the core of a very different transistor model. \ldots`

2 断行和断页

2.1 合理分段

- `\` or `\newline`:

断行但不是另起一段。`\`也在表格、公式等地方用于分行，而 `\newline` 只用于文本段落中。

- `\`*:

断行，但不另起一页和不断页。

- `\newpage` or `\clearpage`:

通常情况下两个命令都能起到另起一页的作用，但有一些区别：一是在双排版中 `\newpage` 只起到另起一栏的作用；二是涉及到浮动体的排版上行为不同。

- `\linebreak[⟨n⟩]` `\nolinebreak[⟨n⟩]` `\pagebreak[⟨n⟩]` `\nopagebreak[⟨n⟩]`:

不满足于 \LaTeX 默认的断行和断页位置，用其高速哪些地方适合断页，哪些地方不合适。 $\langle n \rangle$ 代表合适/不合适的程度，取值范围为 0-4，不带可选参数时，缺省为 4。以上命令适合给出优先考虑断行断页/禁止断行断页的位置，但不适合直接拿来断行或断页，使用 `\newline` 或 `\newpage` 等是更好的选择。

2.2 连字符

如果遇到很长的英文单词，仅在单词之间的位置断行无法生成宽度匀称的行时，就要考虑从单词中间段开。对于绝大部分单词， \LaTeX 能够找到合适的断词位置，在断开的行尾加上连字符 -。如果一些单词没能自动断词，我们可以在单词内手动使用 `\-` 命令指定断词的位置。另外，也可以使用 `\hyphenation{word list}` 命令来指定使用连字符的位置，例如 `texttt\hyphenation{FORTRAN Hy-phen-a-tion}`，其中的 `word list` 是不区分大小写的。

- `\mbox{text}` 或 `\fbox{text}`:

会避免 `text` 被连字符分开。`\fbox\mbox` 多了个可见的框。

3 预定义好的字符串

- `\today` : 打印当天日期
- `\TeX` : \TeX
- `\LaTeX` : \LaTeX
- `\LaTeXe` : $\text{\LaTeX 2}_{\epsilon}$

4 特殊字符和符号

4.1 引号

- 双引号: “...tex...”

- 单引号: ‘...tex...’
“Please press the ‘x’ key.”

4.2 短划线和连字符

在 \LaTeX 中有下面四种横杠:

- - : - 连字符, 用于连接词语。
daughter-in-law, X-rated
- - - : - 短破折号, 常用于连接数字表示起止范围。
pages 13–67
- - - - : — 长破折号, 常用于表示意思的转换。
yes—or no?
- \$-\$: - 减号
0,1 and −1

4.3 波浪线

- \sim : ~
<http://www.rich.edu/~bush>
- \sim : ~
<http://www.clever.edu/~demo>

4.4 斜杠

- read/write: 不允许用连字符拆分
- read\slash write: 不允许用连字符拆分

4.5 度

- 30°C : 30°C
- 30°C : 30°C
- 86°F : 86°F

4.6 欧元符号

首先需要在导言区加载 `textcomp` 包：

```
\usepackage{textcomp}
```

使用命令输出：

```
\texteuro
```

如果所用的字体不包含欧元符号或者想用别的字体的欧元符号，导入 `eurosym` 宏包，用 `gen` 来替换 `official` 参数可以使用和当前字体匹配的欧元符号：

```
\usepackage[official]{eurosym}
```

- `\texteuro`: €
- `\euro`: €

4.7 省略号

L^AT_EX 提供了命令 `\ldots` 来生成省略号，相对于直接输入三个点的方式更为合理。`\ldots` 和 `\dots` 是两个等效的命令。

-
.....
- `\ldots`:...
- `\dots`:...

4.8 连字

有些相邻的字母在排版时会连接起来，可以通过 `\mbox` 命令避免它们相连。

- `ffshfilfluffia`: ffshfilfluffia
- `f \mbox{ } fshf \mbox{ } ilf \mbox{ } luf \mbox{ } f \mbox{ } ia`

4.9 重音符号和特殊符号

L^AT_EX 支持用命令输入各种西欧语言的特殊符号和重音，重音符号和特殊符号命令列表：

ò	ó	ô	õ
ō	ò	ö	ç
ǒ	ǒ	ó	q
q	o	oo	
œ	Œ	æ	Æ
å	Å		
ø	Ø	l	Ł
l	J	i	ı

5 国际语言支持/中文排版支持

L^AT_EX 对其他很多语言提供了支持。babel 宏包可以用于对各种语言进行适配。排版中文有两种方式，一种是使用 xeCJK 宏包，另一种是使用 C_T_EX 宏包和文档类。C_T_EX 宏包和文档类是对 CJK 和 xeCJK 等宏包的进一步封装。文档类包括 ctexart、ctexrep、ctexbook，分别是对 L^AT_EX 的三个标准文档类 article、report、book 的封装，对 L^AT_EX 的排版样式做了许多调整，以切合中文排版风格。最新版本的 C_T_EX 宏包/文档类甚至支持自动配置字体。

5.1 C_T_EX 的安装

C_T_EX 宏集依赖的宏包和宏集已被最常见的 T_EX 发行版 T_EXLive 和 MiK_T_EX 所收录。如果本地安装的 T_EXLive 或 MiK_T_EX 不是完全版本，就需要通过这两个发行版提供的宏包管理器来安装宏包。

T_EXLive 的宏包管理器是 tlmgr。在 Linux 系统上，一般需要 sudo 权限才能正确地执行 tlmgr 的功能。

直接使用 sudo tlmgr [arg] 时，可能会提示找不到 tlmgr 或没有这个命令。sudo 有一种内置的保护机制，只会使用安全的环境变量 PATH。如果 T_EXLive 的路径不在 sudo 的安全环境变量内，它就找不到相关的命令。可以在终端执行 sudo gedit /etc/sudoers，然后将 T_EXLive 的路径添加到 sudo 的 secure_path 中。不同路径用：隔开。

问题解决后，在终端中依次执行以下命令，以更新 tlmgr 宏包管理器、已安装的所有宏包、安装 C_T_EX 宏集。

```
sudo tlmgr update --self
sudo tlmgr update --all
sudo tlmgr install ctex
```

5.2 C_T_EX 文档类

C_T_EX 宏集提供了四个中文文档类：ctexart、ctexrep、ctexbook 和 ctexbeamer，分别对应 L^AT_EX 的标准文档类 article、report、book 和 beamer。使用它们的时候，需要将涉及到的所有源文件使用 UTF-8 编码保存。

下面是使用 ctexart 文档类编写的一个例子：

```
\documentclass[UTF8]{ctexart}
\begin{document}
\end{document}
```

6 单词间的空格

L^AT_EX 默认句子以句点、问号或者感叹号结尾。但是如果句点跟在一个大写字母后面，它不会认为这是句子结尾，因为大写字母后面跟句点往往是缩略词。

用户可以通过具体的命令来改变上面的默认设定。一个斜杠跟一个空格会产生一个不会被扩大的空格；一个波浪线（~）会产生一个既不能被扩大、也不能从这里断行的空格；在句点前使用 \@ 命令，不管这个句点是不是跟在大写字母后面，都会指定这个句子到句点就结束。使用 \frenchspacing 命令可以强制不在一个句子后面插入多余的空格。如果使用 \frenchspacing 命令就没必要再用 \@ 了。

例子：

Mr. Smith was happy to see her

cf. Fig. 5

I like BASIC. What about you?

7 标题、章、节

文档类中的几种分层次结构命令：

`\section{...}`

`\subsection{...}`

`\subsubsection{...}`

`\paragraph{...}`

`\subparagraph{...}`

如果你想把你的文件分成不同的部分而不影响章节编号的使用：

`\part{...}`

当使用 `report` 和 `book` 类时：

`\chapter{...}`

下面是两个比较特殊的情况：

- `\part` 命令不会影响 `chapter` 或 `section` 的编号。
- `\appendix` 命令没有任何参数，会把 `chapter`（对于 `report`、`book`）或 `section`（对于 `article`）的数字编号转换成字母编号。

`\tableofcontents` 命令可以用于建立目录，目录就会在这条命令所在的位置生成。一般新写的文档需要编译两次才能正确生成目录，必要的时候 \LaTeX 也会提示需要编译三次。

以上的分章节的命令都有一个可以在命令名称后加一个星号的版本，例如 `\section{...}` 命令，加星号之后的命令为 `\section*{...}`，不同之处是加星版本的章节命令对应的标题不会显示在目录里，也不会被编号。

有时候章节标题太长，这回导致其在目录里显示不佳。可以通过下面的命令在真真的标题前选择添加一个参数，指定在目录中显示的标题。

`\chapter[Title for the table of contents]{A long and espexially boring title, show in the text}`

整个文档的标题是通过 `\maketitle` 命令产生。在调用 `\maketitle` 命令之前文档标题的内容需要由 `\title{...}`、`\author{...}`、`\date{...}`（可选）等参数指定。在 `\author{...}` 命令的参数中，可以用 `\and` 来间隔多个作者名字。

\LaTeX 2_ϵ 在 `book` 文档有以下三个额外的命令，可以进行前沿、正文、后记的结构划分。这三个命令还可以和 `\appendix` 命令结合，生成有前沿、正文、附录、后记四部分的文档。

- `\frontmatter` 前言部分，放置在文档主体的最开始（`\begin*{document}`），他会把页码变成罗马数字，其后的 `\chapter` 不编号
- `\mainmatter` 正文部分，页码为阿拉伯数字格式，从 1 开始计数，其后的章节编号正常

- `\backmatter` 后记部分，页码格式不变，继续正常计数；其后的`\chapter` 不编号

8 交叉引用

当需要对图片表格等进行引用时，用下列命令：

- `\label{marker}`
- `\ref{marker}`
- `\pageref{marker}`

其中，`marker` 是由用户自行定义的标识符。

A reference to this subsection look like: “see section 8 on page 7.”

9 脚注

可以使用`\footnote{...}` 命令来添加脚注，脚注应该紧跟在它注解的词或句子（包括标点符号）后面。由于脚注会分散读者的注意力，所以尽量在文章主体说清楚，少用脚注。

Footnotes¹ are often used by people using \LaTeX .

10 强调

在 \LaTeX 中可以通过`\underline{...}` 命令来实现。但在印刷书籍中，一般通过 `\emph{...}` 命令，使用意大利字体进行强调。但并不是绝对的，需要结合具体语境。

If you use emphasized text, then \LaTeX uses the normal font for emphasizing.

¹This is a footnote.