

# X<sub>Y</sub>TEX / L<sup>A</sup>T<sub>E</sub>X 中文排版之胡言乱语

作者: lyanry

Email: [lyanry@gmail.com](mailto:lyanry@gmail.com)

## § 1 X<sub>Y</sub>TEX 文稿基本格式

文稿（即用于排版的源文件）包含两部分内容：一部分是正文，也就是需要排版输出的内容；另一部分是排版控制命令，用于控制版面式样、字体、字形等格式。TEX 文稿通常以 `tex` 为文件扩展名。

排版控制命令是以反斜线“\”开头的字串。有一些排版控制命令带有一些参数，由参数来修改其默认行为。排版控制命令的参数有些属于可省略的，有些属于不可省略的。在排版控制命令中，可省略的参数(若不提供这些参数，L<sup>A</sup>T<sub>E</sub>X 采用默认参数)置于方括号中，不可省略的参数(必须要提供的参数)置于花括号中。具体格式可表示如下：

```
\命令名[可省略的参数]{不可省略的参数}
```

X<sub>Y</sub>TEX / L<sup>A</sup>T<sub>E</sub>X 的纯西文的文稿基本格式如下：

```
\documentclass[11pt,a4paper]{article}
\begin{document}
Hello World!
\end{document}
```

上面的文稿中，排版控制命令 `\documentclass` 的可忽略参数告诉 L<sup>A</sup>T<sub>E</sub>X 系统，用户使用的是 A4 纸 (`a4paper`)，正文字体为 11pt，接近中文五号字；不可忽略参数告诉 L<sup>A</sup>T<sub>E</sub>X，用户要撰写一篇论文，这样 L<sup>A</sup>T<sub>E</sub>X 系统便会为用户准备好论文排版的默认环境。除了论文类别，L<sup>A</sup>T<sub>E</sub>X 还提供了书籍 (`book`)、书信 (`letter`)、报告 (`report`) 等。排版控制命令 `\begin{document}` 与 `\end{document}` 表示文稿内容的起始与终止。在 `\documentclass` 与 `\begin{document}` 之间的区域称为导言区，可在此区域内放置一些可影响文档整体排版样式的控制命令。

X<sub>Y</sub>TEX / L<sup>A</sup>T<sub>E</sub>X 的中文文稿与西文文稿没什么区别，仅仅是文稿内容中使用的是中文，如下：

```
\documentclass[11pt,a4paper]{article}
\begin{document}
世界，你好！
\end{document}
```

## § 2 X<sub>Y</sub>TeX 中文文档处理

对 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 可使用 `xelatex` 命令处理生成 pdf 文档:

```
$ xelatex filename.tex
```

现在, 假定上一节中组为示例所列举的中文 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 文稿的文件名为 `example.tex`, 使用 `xelatex` 命令处理该文稿可以生成 `example.pdf` 文档, 但是使用 pdf 阅读器打开 `example.pdf`, 就会发现这是一个空白文档, 而没有如我们所预期的那样会在文档中显示出“世界, 你好!”这是因为 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 并没有为中文文稿指定默认字体, 这需要我们自行设定。这也意味着一个很重要的问题: X<sub>Y</sub>TeX 项目解决了 TeX 国际化的问题, 而我们要解决 X<sub>Y</sub>TeX 本地化问题。但是目前, 国内对 X<sub>Y</sub>TeX 很了解的人太少了, 还未有人提出通用的 X<sub>Y</sub>TeX 中文解决方案, 因此要使用 X<sub>Y</sub>TeX 排出符合中文习惯的文章, 就需要熟悉一些 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 宏包与排版控制命令。

宏包 `fontspec` 可与 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 配合使用可实现在 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 文稿中使用系统自带字体的功能。在 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 文稿中的导言区, 使用 `\usepackage` 指令可加载指定宏包。加载 `fontspec` 宏包后, 使用其提供的 `\setmainfont` 命令可设定文稿正文中的中文字体。对上一节中的中文 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 文稿 `example.tex` 修改如下:

```
\documentclass[11pt,a4paper]{article}
\usepackage{fontspec}
\setmainfont{Adobe Song Std}
\begin{document}
世界,你好!
\end{document}
```

字体设置命令 `\setmainfont` 将 `Adobe Song Std` 指定为文档正文默认字体。Adobe Song Std 是 Adobe 发布 Adobe Reader 8.0 时附带的一款中文宋体, 另外还有一款中文黑体 `Adobe HeiTi Std`, 它们都是免费字体, 可以自由使用。如果你没有装这两款字体, 可以使用 `fc-list` 命令查看系统已安装的字体名录, 如下:

```
$ fc-list :lang=zh-cn
文鼎 P L 简报宋,AR PL SungtiL GB:style=Regular
文鼎PL中楷Uni,AR PL ZenKai Uni:style=Medium
... ..
```

将 `fc-list` 输出结果中的字体名填到 `\setmainfont` 命令中, 即可使得 X<sub>Y</sub>TeX / L<sup>A</sup>TeX 在系统字体目录下找到相应字体并将其嵌入到所生成的 pdf 文档中。虽然可以将 Windows 中文字体挪到 Linux 下使用, 但是现在许多自由抑或免费的中文字体已经可以满足中文排版需要了, 因此, 我们应当尽量不要再去那些私权字体。

现在，使用 `xelatex` 对修改后的 `example.tex` 进行处理，可以生成以中文五号宋体显示“世界，你好！”的单个 pdf 文档。

下面，继续进行中文字体的设置，对 `example.tex` 修改如下：

```
\documentclass[11pt,a4paper]{article}
\usepackage{fontspec}

\setmainfont[BoldFont=Adobe Heiti Std]{Adobe Song Std}
\setsansfont[BoldFont=Adobe Heiti Std]{AR PL KaitiM GB}
\setmonofont{Bitstream Vera Sans Mono}

\begin{document}
世界,你好!
\end{document}
```

在解释修改后的 `example.tex` 所发生的变化之前，我们应当了解一下有关字体的一些常识。

西方国家的字母体系可分为两大字族 (Font Family)：Serif 与 Sans Serif。除此之外，还有一种打印机字体虽然也是 Sans Serif，但由于它是等距字，所以又独立出一个 Typewriter 字族。

Serif，中文常译为“衬线”，Sans Serif 则译为“无衬线”。衬线字体是源于古代在一些岩石或金属上刻字时，雕刻刀在笔画的起落处要有入刀与退刀的讲究，不然会损伤刻刀。无衬线字体，是相对于衬线字体而言的。对于西文的衬线字体与无衬线字体的直观意象可见：

N N

中文的字族可分为：隶、楷、行、宋、仿宋、黑、幼圆等，要与西文字族相对应(计算机是西方文明的产物，汗)，那么宋、仿宋都可以看作是衬线字体，而楷体、黑体、幼圆可以看作是非衬线字体。

一旦理解了这些字体常识，对于 `example.tex` 中新增加的那几条设置中文字体的控制命令应该明白个三五分。譬如，`\setsansfont` 指令是设定无衬线中文字体的，我们在 `example.tex` 中使用该指令将无衬线字体设置为楷体 AR PL KaitiM GB。`\setmainfont` 是设置衬线字体的，因为 X<sub>Y</sub>TeX 将衬线字体视为文档默认字体族，而 X<sub>Y</sub>TeX 之所以如此，是因为在实践有一个结论：衬线字体作为文章的正文字体可使读者长时间阅读文章视觉不疲倦。非衬线字体在文章中适合作为标题出现，因为它较衬线字体更为醒目，但如果用无衬线字体作为文章的正文字体，长时间阅读，很容易出现视觉疲劳。`fontspec` 宏包还提供了与 `\setmainfont` 等价的命令 `\setromanfont`，这完全是出于历史的缘故，因为 Roman 字体在西方一向被认识是文章正文字体的正统，最有名的是 Times New Roman。

下面讲一下 `\setmainfont` 与 `\setsansfont` 指令中的可省略参数 `BoldFont` 的用法，这个参数是用来指定衬线与非衬线字体在粗体 (`bold`) 状态下所使用的字体，这是因为字体可以在常态下经“加粗”后所得到的实际上是另一种字体。对于任意一款计算机字体而言，它不是一个你想怎么变就可以怎么变的东西，如果一款字体在设计的时候就不是粗体，那么是不可能把它变成粗体的，只有用一种设计好的粗体去替换。虽然有一些办法可以让一些字体经过微量平移并叠合后可以得到类似“粗体”的效果，但那是“穷人的粗体”，显示效果很差的。所以，我们不应该把你正在用的这个“宋体”变成“粗”宋体，而必须去找专门的粗宋体来用。如果找不到粗宋体，那就用黑体来代替，本文的排版就是这么做的，采用 `Adobe Heiti Std` 来作为衬线与非衬线的粗体。

字体设置完成后，就要考虑中文断行的问题。还是那句话，`XeTeX` 只致力于解决国际化问题，并不考虑本地化，如果说考虑了，那也只是默认考虑了西文本地化。西文的断行问题是根据单词之间的空格来决定一行文本中在哪个单词的尾部断开产生新行的，对中文而言，这种方法就不适用了，因为中文不是以空格来划分单词的。但不要以为 `XeTeX` 不能很好的处理中文断行，在 `XeTeX` 内部已有人为中文断行写了一些规则，我们可以直接使用它们，可在 `XeTeX` / `LaTeX` 的导言区中添加以下指令：

```
\XeTeXlinebreaklocale "zh"  
\XeTeXlinebreakskip = 0pt plus 1pt minus 0.1pt
```

上述指令中，`\XeTeXlinebreaklocale` 指定使用中文断行规则，`\XeTeXlinebreakskip` 可以让 `XeTeX` 处理中文断行时多一点点自适应调整的空间。

Okay! 事实上，讲到这里，`XeTeX` / `LaTeX` 的用法基本已讲述完毕，剩下内容就是 `TeX` / `LaTeX` 的使用了，它们的许多教程都基本适用 `XeTeX` / `LaTeX`，只是有关 `CJK` 那部分不适用。

## § 3 论文版式

一篇论文应该包括两个层次的含义：内容与表现，前者是指文章作者用来表达自己思想的文字、图片、表格、公式及整个文章的章节段落结构等，后者则是指论文页面大小、边距、各种字体、字号等。一篇排版良好的论文应当是内容与表现分离的。本节主要介绍如何使用  $\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  定义论文的表现。

### § 3.1 准备纸张

首先准备纸张，在 `\documentclass` 的可省略参数中，A4 纸用 `a4paper` 表示，A5 纸用 `a5paper`，其他型号用纸的表示类推便是。如果在 `\documentclass` 指令中未指定纸张型号，则  $\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  默认用纸是美国信纸 ( $14 \times 8.5in$ )。

纸张准备好了，然后就是设置基本字体尺寸。一般而言，中文小四号字用像素点为单位表示为 `12pt`，中文五号字表示为 `11pt`。基本字体尺寸的设定非常重要，譬如行距、段落缩进、页芯等参数， $\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  会基于基本字体尺寸给出相应的默认值。基本字体尺寸也是在 `\documentclass` 指令中作为其可省略参数进行设定的，如果未设定该参数，则  $\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  会以 `10pt` 为默认值。

现在，若要在一张 A4 纸上以 `11pt` 为基本字体尺寸写一篇论文，`\documentclass` 指令可写为：

```
\documentclass[a4paper,11pt]{article}
```

$\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  默认是纵向模式排版，要改为横向排版，可添加 `\documentclass` 命令的可省略参数 `landscape`：

```
\documentclass[a4paper,11pt,landscape]{article}
```

`\documentclass` 还有一些常用的可省略参数，比如 `titlepage` 可以让文章的标题单独占据一页，`notitlepage` 可使标题与文章正文排在同一页面。又比如 `draft` 可以控制  $\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  在超出页面宽度限制的文本行右端显示一个粗黑条，提醒用户注意，而 `final` 的作用恰好相反，无论文本行超出边界多少，也不显示粗黑条，但  $\text{X}\_{\text{T}}\text{E}\_{\text{X}}/\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  在编译  $\text{T}\_{\text{E}}\text{X}$  文档时，会给出警告。

### § 3.2 设置页面边距

下面谈谈页边距的设置。MS Word 默认的页面边距为：

- 上边距=下边距=  $1in$  ( $2.54cm$ )
- 左边距=右边距=  $1.25in$  ( $3.17cm$ )

使用宏包 `geometry` 可以进行  $\text{X}\_{\text{Y}}\text{T}\_{\text{E}}\text{X}$  /  $\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  文稿的页面边距设置：

```
\usepackage[top=1in,bottom=1in,left=1.25in,right=1.25in]{geometry}
```

实际上这样设置的页面边距极不美观，尤其是左右对称的页边距没有考虑装订的需要，另外上边距如果加上页眉或就显得过窄。因此，要是真的很注重页面美观的话还是自己去调整一下，比如我喜欢将页边距设置下面这样：

```
\usepackage[top=1.2in,bottom=1.2in,left=1.2in,right=1in]{geometry}
```

将左边距设置的比右边距大一些，主要是考虑装订的需要，但是在实际打印时有单面打印与双面打印模式，在双面打印时，应该是奇数页面的左边距比右边距大一些，在偶数页则相反。 $\text{X}\_{\text{Y}}\text{T}\_{\text{E}}\text{X}$  /  $\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  考虑到了这一点，在偶数页面中会自动将左、右边距切换。指定文稿单双页面的参数有 `oneside` 与 `twoside`，它们都是 `\documentclass` 的可省略参数，如果文稿类别是论文，默认是单面打印模式。

如果相对 `geometry` 宏包的使用进行更详细的了解，请参考文献 [?]

### § 3.3 章节标题

可使用 `titlesec` 宏包设置章节标题。在引入 `titlesec` 宏包时，可以指定一些格式选项，比如：

```
\usepackage[center,pagestyles]{titlesec}
```

其中 `center` 可使标题居中，还可设为 `raggedleft` (居左，默认)，`raggedright` (居右)。`pagestyles` 是申明后面要使用 `titlesec` 宏包自定义页面样式 (在下一节会讲)。

标题由标签+标题内容构成，其格式通常在  $\text{X}\_{\text{Y}}\text{T}\_{\text{E}}\text{X}$  /  $\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  文稿的导言区中设置。要设置论文中的节标题格式，可用 `titleformat` 指令，用法如下：

```
\titleformat{command}[shape]{format}{label}{sep}{before}[after]
```

其中各参数含义如下：

- `command` 是要重新定义的各种标题命令，比如 `\section`、`\subsection`，还有更多的，在后文中讲书籍排版时再谈；
- `shape` 是用来设定段落形状的，可选的参数有 `hang`、`block`、`display` 等，详见 `titlesec` 文档，位于：  
`$\text{TEXLIVE}/\text{\$VERSION}/\text{texmf-dist}/\text{doc}/\text{latex}/\text{titlesec}`
- `format` 用于定义标题外观，比如使标题居中、字体加粗等；

- `label` 用于定义定义标题的标签，就是标题内容前面的标号；
- `sep` 定义标题的标签与标题内容之间的间隔距离；
- `before` 用于在标题内容前再加些内容；
- `after` 用于在标题内容后再加些内容；

本文排版所用节标题分为两级，其格式采用以下命令设置：

```
\titleformat{\section}{\centering\Large\bfseries}{\S,\thesection}{1em}{}
\titleformat{\subsection}{\large\bfseries}{\S,\thesubsection}{1em}{}
```

其中，`shape`、`before`、`after` 参数都被省略掉了。`format` 参数将 `section` 格式设置为居中（`\centering`）、字号为 `\Large`，字体被加粗显示 `\bfseries`；在设置 `subsection` 格式，未采用居中，而是采用默认的居左，另外将标题的字号也降了一级（`\large`）。`label` 参数将标题的标签设置为以“§”为前缀 + 标题序号。`sep` 参数设置标签与标题内容之间以一个字(1em)的宽度为间隔。

### § 3.4 页眉与页脚

这一节讲怎样使用 `titlesec` 宏包设置页眉、页脚。下面的命令在 `XYLATEX / LATEX` 导言区定义了一个新的页面样式，并使用该样式：

```
\newpagestyle{main}{
  \sethead{\small\S,\thesection\quad\sectiontitle}{\S\cdots~\thepage~\cdots}
  \setfoot{}{}\headrule}
\pagestyle{main}
```

其中 `\sethead` 命令设置页眉，用法为：

```
\sethead[偶数页左页眉][偶数页中页眉][偶数页右页眉]
{奇数页左页眉}{奇数页中页眉}{奇数页右页眉}
```

单面打印模式只要给出奇数页的设置即可，双面模式则需要将左、右页眉做个调换。上面给出的例子是单面模式的。`\setfoot` 指令用法与 `\sethead` 用法相似。

上面的页眉页脚设置示例中，`\headrule` 指令可画页眉线，默认宽度是 0.4pt，如果对该宽度不满意，可使用下面命令重新设置其宽度：

```
\setheadrule{宽度值}
```

上面的页眉设置示例的排版效果即本文档页眉效果。



### § 3.5 小结

以上讲述了有关论文文稿的纸张准备、页边距设置、章节标题以及页眉页脚设置等内容，其中有些  $\text{X}\_{\text{Y}}\text{T}\_{\text{E}}\text{X}$  /  $\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  命令对于初学者而言，也许暂时不能领会其作用，这没有关系，可以先建立一个大致意象，待日后对  $\text{X}\_{\text{Y}}\text{T}\_{\text{E}}\text{X}$  /  $\text{L}\_{\text{A}}\text{T}\_{\text{E}}\text{X}$  语法较为熟悉后再回头来看就可以了。另外，对于论文排版，还有许多排版元素的设置这里没有提到，比如文章标题、摘要与参考文献等，这些排版元素我将其划分为论文的内容部分，会在后文中的论文与书籍的内容排版中详细讲述。

由于要赶一份材料，本文档暂时进行到这里，待那份材料完工后再来写。