

中文 T_EX 简介

张卫斌

2018 年 6 月 5 日

目录

第一章 简介	2
1.1 文件类	2
1.1.1 安装包	2
1.1.2 字体	2
1.2 编译	2
第二章 字体	3
2.1 字样	3
2.2 字体格式	4
2.2.1 点阵字体和矢量字体	4
2.2.2 常见字体	4
2.2.3 合纵连横	5
2.3 字体应用	6
2.3.1 DVI	6
2.3.2 dvips	6
2.3.3 dvipdfm(x)	6
2.4 TrueType 字体安装配置	7
2.4.1 目录和文件	7
2.4.2 ttf2tfm	8
2.4.3 字体定义文件	8
2.4.4 配置 ttf2pk	8
2.4.5 配置 dvipdfmx	9

第一章 简介

你好吗？

1.1 文件类

可用的文件类有: ctexart, ctexrep, ctexbook. 即只需在原类的基础上加上 ctex, 同时缩写类名即可。

注意: 并没有 ctexletter 类型。

1.1.1 安装包

为了适用 ctexart, ctexrep, ctexbook 等文件类, 我们需要安装 `texlive-lang-chinese` 包以安装相关的类文件: ctexart.cls ...

1.1.2 字体

中文字体可用文泉驿字体 (<http://wenq.org/wqy2/index.cgi?Home>), 下载选中字体, 解压缩后将字体文件 (*.ttc) 放置在 /usr/share/fonts/wen-quanyi 目录下即可。

1.2 编译

中文 L^AT_EX 文档需用 xelatex 编译。

第二章 字体

关于字体有三个重要概念：glyph、typeface、font。glyph 通常被翻译为字形，也有翻译为字体的；typeface 是一个书法和印刷领域的概念，它通常被翻译为字体或书体；font 曾经和 typeface 混用，但现在一般用作电脑领域的概念，在中国大陆被翻译为字体，在台湾被翻译为字型。

上述翻译的混乱令人十分无奈，包老师决定在本文中把它们分别翻译为字形、字样、字体，以正视听。

字形是一个字符的具体图形表现形式，一个字符可以有多个字形，比如汉字中的“強 / 强”、“戶 / 户 / 戸”；字样是一组相同风格样式的字形的集合，比如中文字样有宋、仿、楷、黑、隶、篆等；一种字样可以对应电脑上的几种字体。

2.1 字样

拉丁字母的字样主要有三大类：Serif (Roman)、Sans Serif 和 Monospace (Typewriter)。Serif 的笔画边缘部分有些装饰，类似于中文的宋体、仿宋、楷体、魏书等。Sans Serif 的笔画则是平滑的，类似于中文的黑体。Sans 这个词来源于法语，就是“没有”的意思。Monospace 则是等宽字样。

每一类字样都可以有加粗 (bold)、斜体 (italic)、倾斜 (oblique) 等修饰效果。Italic 通常对原字样进行过重新设计，它修饰精细，多用于 Serif；Oblique 也称作 slanted，基本上是把正体倾斜，多用于 Sans Serif。通常 oblique 看起来比 italic 要宽一些。

2.1列出了几种常见的字样。

表 2.1: 常见字样

操作系统	Serif	Sans Serif	Monospace
Mac OS	Times	Helvetica	Courier
Windows	Times New Roman	Arial	Courier New

2.2 字体格式

2.2.1 点阵字体和矢量字体

电脑上用的字体 (font) 按数据格式可以分为三大类: 点阵字体 (bitmap)、轮廓 (outline) 字体和笔画 (stroke-based) 字体。

点阵字体通过点阵来描述字形。早期的电脑受到容量和绘图速度的限制, 多采用点阵字体。点阵字体后来渐渐被轮廓字体所取代, 但是很多小字号字体仍然使用它, 因为这种情况下轮廓字体缩放太多会导致笔画不清晰。

轮廓字体又称作矢量字体, 它通过一组直线段和曲线来描述字形。轮廓字体易于通过数学函数进行缩放等变换, 形成平滑的轮廓。轮廓字体的主要缺陷在于它所采用的贝塞尔曲线 (Bézier curves) 在光栅 (raster) 设备 (比如显示器和打印机) 上不能精确渲染, 因而需要额外的补偿处理比如字体微调 (font hinting)。但是随着电脑硬件的发展, 人们一般不在意它比点阵字体多出的处理时间。

笔画字体其实也是轮廓字体, 不过它描述的不是完整的字形, 而是笔画。它多用于东亚文字。

2.2.2 常见字体

常见的轮廓字体技术有: Type 1 和 Type 3、TrueType、OpenType、MetaFont 等。

Adobe 的 Type 1 和 Type 3 基于 PS, 它们采用三次贝塞尔曲线。Type 1 支持微调, 它使用一个简化的 PS 子集; Type 3 不支持微调, 但它可以使用全部 PS 功能, 因此既可以包含轮廓字体也可以包含点阵字体信息。

1991 年, Apple 发布了 TrueType, 它采用二次贝塞尔曲线。二次曲线处理起来比三次曲线快, 但是需要更多的点来描述。所以从 TrueType 到 Type 1 的转换是无损的, 反之是有损的。1994 年, Apple 着手研究 TrueType 的

下一代技术：TrueType GX，它后来演变为 Apple Advanced Typography (AAT)。

1996 年，微软和 Adobe 联合发布了 OpenType。它比起 AAT 的优势有：跨平台、开放和易于开发、支持更多的语言比如阿拉伯语。

早在 1984 年 Knuth 就发布了 MetaFont，它与 TrueType 和 OpenType 的区别是，不直接描述字形轮廓，而描述生成轮廓的笔的轨迹。笔的形状可以是椭圆形或多边形，尺寸缩放自如，字形边缘也柔和一些。两种字体可以用同一个 MetaFont 文件，当然还有不同的参数。MetaFont 技术如此先进，却没有流行开来。对此 Knuth 解释道，要求一位设计字体的艺术家掌握 60 个参数太变态了，那是用来折磨数学家的。

Type 1 和 Type 3 把字体信息存储在两种文件里：metrics 和 glyph 文件。metrics 文件有 AFM (Adobe font metrics) 和 PFM (printer font metrics)，glyph 文件有 PFA (printer font ASCII) 和 PFB (printer font binary)。L^AT_EX 使用的 metrics 格式是 TFM (TeX Font Metrics)。

TrueType 的文件后缀是 .ttf，OpenType 的是 .ttf 和 .otf。MetaFont 虽然用矢量图形来定义字形，实际输出的却是一种点阵格式：PK (packed raster)。

上述字体按技术的先进性，从高到低的排序为：OpenType、TrueType、Type 1、Type 3、PK，我们应优先选用 OpenType 和 TrueType。

2.2.3 合纵连横

Adobe 收取的 Type 1 专利许可费一度十分昂贵，穷人们只好用免费的 Type 3。为了打破这种垄断，Apple 开发了 TrueType。1991 年 TrueType 发布之后，Adobe 随即公开了 Type 1 的规范，Type 1 字体从贵族堕落为平民，因而流行开来。

1980 年代中后期，Adobe 的大部分盈利来自于 PS 解释器的许可费。面对这种垄断局面，微软和 Apple 联合了起来。微软把买来的 PS 解释器 TrueImage 授权给 Apple，Apple 则把 TrueType 授权给微软。

微软得陇望蜀，又企图获得 AAT 的许可证，未遂。为了打破 Apple 的垄断，微软联合 Adobe 在 1996 年发布了 OpenType。Adobe 在 2002 年末将其字体库全面转向 OpenType。

上面这几出精彩好戏充分展示了商场上的勾心斗角、尔虞我诈，没有永恒的伙伴，只有永恒的利益。但它同时也告诉我们，市场竞争中受益的还是

广大的消费者。

2.3 字体应用

PS 支持 Type 1 和 Type 3, 而 PDF 除了这两种还支持 TrueType 和 OpenType。L^AT_EX、DVI 浏览器、各种 driver 分别采用不同的字体技术。

2.3.1 DVI

latex 编译 L^AT_EX 源文件生成 DVI 时只需要 .tfm 文件, 因为 DVI 并不包含字形信息, 而只包含对字体的引用。DVI 浏览器显示 DVI 时一般使用 PK, 它在系统中查找相应的 .pk 文件, 若找不到就调用 MetaFont 在后台自动生成。

2.3.2 dvips

缺省情况下, dvips 也会查找 .pk, 或调用 MetaFont 自动生成; 然后把 PK 转换成包含点阵字体的 Type 3, 它的参数 -D 可以用来控制该点阵字体的分辨率。用 ps2pdf 处理含 Type 3 的 PS 时, 输出的自然是含 Type 3 的 PDF。

GSview 在低分辨率下可以很好地渲染 Type 3, Adobe Reader 或 Acrobat 却不能, 因为它们使用的 Adobe Type Manager 不支持包含完整 PS 的 Type 3。含 Type 3 的 PDF 看起来会有些模糊, 所以应尽量避免使用。

dvips 的另一个参数 -Ppdf 把 Type 1 嵌入生成的 PS, 这样再 ps2pdf 就能生成含 Type 1 的 PDF。

dvips 不支持真正的 (native) TrueType, 用户只能把 TrueType 先转换成 PK 或 Type 1, 这样绕了个弯效果总会打些折扣。

dvips 的字体详细使用方法可查阅其手册 [?] 第 6 章, 此处不赘述。

2.3.3 dvipdfm(x)

dvipdfm 支持 PK 和 Type 1, 它可以用一个 t1fonts.map 文件建立 PK 文件和 Type 1 文件之间的映射, 这样生成的 PDF 用的就是 Type 1。dvipdfm 也不支持真正的 TrueType。

dvipdfmx 通过正确的设置可以使用真正的 TrueType，它对中日韩等东亚文字的支持也较好，所以它对我们来说是 Driver 的首选。

2.4 TrueType 字体安装配置

CJK 自带的 UTF-8 编码字体 gbsn 和 gkai 只包含 GB2312 字符集，而 CTeX 只提供 GBK 编码字体，因此中文用户通常需要自己安装配置 UTF-8 编码的 TrueType 字体。

在使用 TrueType 之前，用户通常需要作以下准备工作：

1. 用转换程序 `ttf2tfm` 生成 TFM。
2. 配置字体定义文件 `.fd`。
3. 配置 `ttf2pk`，因为 DVI 浏览器和 `dvips` 都会自动调用 `ttf2pk` 来生成 PK。
4. 配置 `dvipdfmx`。

2.4.1 目录和文件

通常每个发行包都会参照 TDS 建立自己的目录系统，把各种文件发在固定的位置。比如 MiKTeX 顶层目录如下，在本节后面的示例中我们将使用这些目录的缩写。

```
Install: D:\edit\MiKTeX 2.7
UserData: C:\Documents and Settings\Alpha\Local Settings\
Application Data\MiKTeX\2.7
UserConfig: C:\Documents and Settings\Alpha\
Application Data\MiKTeX\2.7
```

目录多了有个缺点，文件不知道放在哪里好。MiKTeX 中有的配置文件居然在四个目录下各有一份，实在是令人发指。幸好我们可以用下面的命令检查配置文件的具体名字和路径。

```
initexmf --edit-config-file=ttf2pk
```


2.4.2 ttf2tfm

比如我们想把 `SimSun18030.ttc` (18030 字符集的新宋体) 转换为 UTF8 编码的字体文件, 我们需要执行以下步骤。

1. 把需要的 `.ttf` 文件复制到 `UserData/fonts/truetype/chinese/`。
2. 用下面的命令生成 `.tfm` 和 `.enc` 文件。
3. 把 `*.tfm` 复制到 `UserData/fonts/tfm/chinese/utf8song/`。
4. 把 `*.enc` 复制到 `UserData/fonts/enc/chinese/utf8song/`。

```
ttf2tfm SimSun18030.ttc -q -w utf8song@Unicode@
```

2.4.3 字体定义文件

字体定义文件将字体引用名和实际的字体文件联系起来, 比如我们在 CJK 环境中引用 `usong` 时, 系统将会找到并使用 `utf8song*.tfm`。

```
%UserData\tex\latex\CJK\UTF8\C70usong.fd
\ProvidesFile{c70usong.fd}
%character set: GB18030
%font encoding: Unicode
\DeclareFontFamily{C70}{usong}{\hyphenchar \font\m@ne}
\DeclareFontShape{C70}{usong}{m}{n}{<-> CJK * utf8song}{}
\DeclareFontShape{C70}{usong}{m}{it}{<-> CJK * utf8song}{}
\DeclareFontShape{C70}{usong}{bx}{n}{<-> CJKb * utf8song}{
  \CJKbold}
\endinput
```

2.4.4 配置 ttf2pk

MiKTeX 中 `ttf2pk` 的配置文件是 `ttf2pk.ini`, 在其它的发行包中可能是 `ttf2pk.cfg`。

`ttf2pk.ini` 中有一个 `.map` 文件列表, 后者定义了 TrueType 应该按编码转为 PK 等信息。

比如下面这个文件列表会让 `ttf2pk` 读取 `foo.map` 和 `bar.map`。

```
map foo.map  
map bar.map
```

如果系统找不到 `ttf2pk.ini`，它会缺省使用 `ttfonts.map`。

```
%UserData\ttf2tfm\base\ttfonts.map  
utf8song@Unicode@ SimSun18030.ttc
```

2.4.5 配置 dvipdfmx

配置 `dvipdfmx` 是为了让 PDF 正确地嵌入 TrueType，否则生成的文件中的内容不能复制、粘贴。

```
%UserConfig\dvipdfm\config\dvipdfmx.cfg  
f cid-x.map
```

```
%UserData\dvipdfm\config\cid-x.map  
utf8song@Unicode@ unicode SimSun18030.ttc
```