

# 关于新版 CCT 的说明

张林波

(zlb@lsec.cc.ac.cn)

中国科学院数学与系统科学研究院

2003 年 2 月 6 日

## §1 构成与功能

新版 CCT 中主要增加了一个宏文件 `CCT.sty`, 同时对老版本中的一些程序和宏文件做了少量修改, 并增加了一些辅助程序和宏文件。新版 CCT 使用新的版本号系列, 目前的版本号是 0.4.0 (老版本目前为 5.14)。

老版本 CCT 的 DOS 版是捆绑  $\text{emT}_{\text{E}}\text{X}$  系统发行的。由于  $\text{emT}_{\text{E}}\text{X}$  不支持高位为 1 的字符, 因此新版 CCT 中的宏文件不能直接在  $\text{emT}_{\text{E}}\text{X}$  中使用, 而需要先进行转换, 参看 §3.3 最后一段中的说明。

新版 CCT 的源文件可从 `ftp://ftp.cc.ac.cn/pub/cct/BETA-TEST/` 下载。

新版 CCT 主要增加了下面一些功能:

1. 排版源文件可以不经预处理程序而直接用  $\text{T}_{\text{E}}\text{X}/\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$  排版, 解决了老版本在 `verbatim` 等环境中不能使用汉字的问题, 也可以方便多个文件的处理。
2. 新版本提供了使用 CJK 字库的选项, 用这种方式可以充分利用许多软件对 CJK 系统的支持, 如使用 Type 1 字库生成矢量的中文 PS 文件, 利用 `dvipdfmx`, `pdflatex` 等程序生成高质量的 PDF 文件, 等等。几乎所有支持 CJK 系统的工具或宏包都可以以某种方式在 CCT 中用。

使用新版 CCT 时, 对一个排版文件的处理可以有两个选择, 即使用预处理程序和不使用预处理程序。如果使用预处理程序, 则排版及后处理过程与老版 CCT 中完全一样, 得到的排版结果也应该一样。如果不使用预处理程序, 则还可以选择生成 CCT 格式的 dvi 文件或生成 CJK 格式的 dvi 文件。如果选择生成 CJK 格式的 dvi 文件, 则 dvi 文件不必进行后处理而可以直接显示或打印, 也可以直接用  $\text{PDF}_{\text{T}_{\text{E}}\text{X}}/\text{PDF}_{\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}}$  生成 PDF 文件。

建议 CCT 排版源文件的命名遵循下述规则: 如果希望使用 CCT 的预处理程序则源文件的扩展名取为 “.ctx”, 否则源文件的扩展名取为 “.tex”。这样在编制一些批处理程序或脚本时可以根据源文件的扩展名来判断是否需要进行预处理, 而根据 .dvi 文件中是否引用了以 `ccfnt` 开头的字库可以判定是否需要运行后处理程序 `patchdvi` (参看 §4 中的 Linux 脚本)。

此外, 由于维护方面的原因, 我们将不再发布新的 CCT DOS 和 Windows 版本。建议 CCT 的 Windows 用户使用  $\text{C}_{\text{T}}\text{E}_{\text{X}}$  (<http://www.ctex.org>) (希望新版 CCT 的功能能尽快被集成到下一个  $\text{C}_{\text{T}}\text{E}_{\text{X}}$  版本中)。

本文档的编译需要 CJK 字库。如果没安装 CJK 字库请将 “\documentclass” 中的 “[CJK]” 选项去掉。

## §2 改动部分

### 1. 预处理程序 `cct.c`

主要改动是当中英文间已经有空格或“~”时不再加入空格。因为 CJK 系统的用户习惯在中英文间加“~”，如“中文~\TeX~...”，新的预处理程序可以避免在这种情况下插入多余的空格。

### 2. 初始化程序 `cctinit.c`

由于 `CCT.sty` 中重新定义了“~”（相当于使用了 CJK 中的 `\CJKtilde` 命令），`cctinit` 中相应修改了 `cchead.sty` 中的一些定义以避免破坏中文行禁则。

### 3. 宏文件 `cctbase.sty`

自动包含 `CCT.sty`，并做了一些相关的修改。

### 4. 文档类 `cctart.cls` 和 `cctbook.cls`

增加了 CJK 选项，用于选择生成 CJK 格式的 dvi 文件。如果用了 CJK 选项，则不能再用 CCT 的预处理程序进行处理（这种情况下排版时会产生警告信息）。

## §3 新增功能

### §3.1 宏文件 `CCT.sty`

这是一个  $\TeX$  宏包，可以替代 CCT 的预处理程序，适用于 Plain  $\TeX$ ， $\LaTeX$ ，以及  $\text{PDF}\TeX$ ， $\text{PDF}\LaTeX$  等。它的主要作用是用  $\TeX$  语言实现了 CCT 预处理程序的功能，从而使得 CCT 排版文件可以不经预处理而直接用  $\TeX$  排版。对用户而言，`CCT.sty` 与 CCT 预处理程序间的主要区别在于 CCT 预处理程序会自动在中英文间加入空格，而 `CCT.sty` 则不会自动插入空格，并且缺省时会吃掉汉字后面的空格，因此不使用 CCT 预处理程序时需要手工在中英文间加空格，并且在汉字后面需要使用“控制空格”（即“\\_”或“~”）以免空格被汉字吃掉。建议用户借鉴 CJK 排版文件的写法，即在中文-英文间加一个“~”，而在英文-中文间加一个空格（可参看本文档的写法）。另一个方法是使用 §3.2 中介绍的辅助程序 `cctspace` 自动在中英文间加空格。

`CCT.sty` 使用一个配置文件 `CCT.cfg` 来定义一些参数，包括各个字号的大小、CCT 字体到 CJK 字体的映射和汉字标点符号。`CCT.cfg` 中定义的字号与字体只有在用户选择了输出 CJK 格式的 dvi 文件时（即在 `\documentclass` 中使用了 `[CJK]` 选项）才起作用。当输出 CCT 格式的 dvi 文件时，字号与字体依然由文件 `cct.dat` 和 `ccfonts.def` 定义。

`CCT.sty` 中还定义了一些用户命令，这些命令的作用与 CJK 中的对应命令类似，包括 `\CCTspace`，`\CCTnospace`，`\CCTtilde`，`\standartilde`，`\nbs`，和 `\CCTchar`。命令 `\CCTchar` 允许用户直接输入汉字编码，如可用 `\CCTchar{176}{161}`（或是 16 进制形式 `\CCTchar{"B0"}{"A1}`）来输入汉字“啊”。该命令通常用于输入一些不易输入的字符。其余命令普通用户一般不用，因此不在此介绍。

`CCT.sty` 的使用方法如下：

#### 1. $\LaTeX$ 2<sub>ε</sub> 系统

如果使用新的中文文档类 `cctart.cls`，`cctbook.cls`（或宏包 `cctbase.sty`），则 `CCT.sty` 已经被自动包含，因此可以直接使用相关的功能，并且可以用可选项“`[CJK]`”来选择生成 CJK 格式的 dvi 文件（没有该可选项时生成 CCT 格式的 dvi 文件）。如果使用其它文档类，则需使用：

```
\usepackage{cctbase}
```

或：

```
\usepackage{ccthead}
```

```
\usepackage{CCT}
```

（“`\usepackage{cctbase}`”或“`\usepackage{CCT}`”中可以使用可选项 `[CJK]` 来选择生成 CJK 格

式的 dvi 文件)。

## 2. 其它系统

需在文档开头加入:

```
\input cthead.sty
\input CCT.sty
```

或:

```
\let\CCTCJKfonts=1
\input cthead.sty
\input CCT.sty
```

(“\let\CCTCJKfonts=1” 选择生成 CJK 格式的 dvi 文件)。

### §3.2 辅助程序 cctspace.c

这是一个辅助程序, 用于自动调整中英文间的空。其用法如下:

```
cctspace 输入文件 >输出文件
```

或:

```
cctspace <输入文件 >输出文件
```

cctspace 可以作为一个辅助工具, 用于对老的 .ctx 文件进行转换以便不用预处理程序处理。也可以看作一个新的“预处理程序”(但它只在输出文件中加入一些空格和 “~”)。

cctspace 运行时会读入一个名为 cctspace.cfg 的配置文件, 其中列出不应该加空格的地方。cctspace.c 按照通常 T<sub>E</sub>X 输入文件的规则搜索 cctspace.cfg 文件, 因此后者可以放在 T<sub>E</sub>X 的某个 input 路径中。

cctspace 偶而会引入不应该有的空格。用户可以通过加入一些列在 cctspace.cfg 中的特殊命令来防止 cctspace 在某些地方加入空格。如“啊ABC”经处理后会变成“啊~ABC”, 如果不希望在“ABC”前面加入空格, 则可写成“啊\relax ABC”的形式, 因为 cctspace 不会在“\relax”的前后加空格(或“~”), 而“\relax”命令本身不起任何作用。

cctspace 程序另外一个有趣的用法是当用户在中英文间手工插入空格时, 用来检查是否有漏加空格的地方。方法是用 cctspace 程序处理排版文件, 然后比较处理前后文件间的差异 (Unix 下可用“diff”命令比较, DOS/Windows 下可用“fc”或“comp”命令比较)。

### §3.3 辅助程序 cctconv.c

GB2312 编码中汉字的两个字节都在 161-254 之间, 用 T<sub>E</sub>X 处理这些汉字不会出问题。GBK 编码是 GB2312 的超集, 在 GBK 编码中, 第一个字节的取值范围为 129-254, 第二个字节的取值范围为 64-254 (不能取 127)。GBK 汉字的第二个字节的取值范围中包含了 T<sub>E</sub>X 特殊字符 ‘\’, ‘{’, ‘}’, ‘^’, ‘\_’, 和 ‘~’, 前四个字符用 T<sub>E</sub>X 处理时会出问题, 而后两个字符在 PDF 书签文件中 (.out 文件) 会出问题, 因此在使用 T<sub>E</sub>X 排版包含这样的汉字的文件时需要做一些特殊处理。辅助程序 cctconv 对汉字中的特殊字符进行转换, 以确保 T<sub>E</sub>X 系统能正确地处理它们 (cctconv 程序将汉字中的字符 ‘\’, ‘{’, ‘}’, ‘^’, ‘\_’, 和 ‘~’ 分别转换成 ‘0’, ‘1’, ‘2’, ‘3’, ‘4’, 和 ‘5’, 再由 CCT.sty 将它们转换回来)。

cctconv 程序的用法如下:

```
cctconv 输入文件 >输出文件
```

或:

```
cctconv <输入文件 >输出文件
```

注意, 如果用户使用 CCT 的预处理程序, 则不能用 `cctconv` 程序对排版文件做处理, 否则包含这些字符的汉字会不对。

`cctconv` 程序有一个命令行可选项 “-f”, 它强制将所有第 8 位为 1 的字符 (即 128–255 范围内的字符) 转换成 “`^^xx`” 的形式, 其中 “xx” 为字符的 16 进制编码。如组成汉字 “啊” 的两个字节的 16 进制编码为 `b0` 和 `a1`, 因此 “`cctconv -f`” 会将 “啊” 转换成 “`^^b0^^a1`”。有些老的  $\text{\TeX}$  系统 (如  $\text{\em\TeX}$ ) 不允许输入文件中直接使用大于 127 的字符, 此时需用 “`cctconv -f`” 对所有含汉字的文件进行转换 (包括用户排版文件、`CCT.cfg`、`cctbase.sty`、`cctart.cls` 等), 然后才能进行处理。

由于 简体中文中很少用到 GB2312 编码之外的汉字, 而且所有汉字中只有数百个含有  $\text{\TeX}$  的特殊字符, 因此 `cctconv` 程序只在极少数特殊情况下有必要。

### §3.4 书签文件转换程序 `gbk2uni.c`

这个程序由 `cxterm` 编制。由于 Acrobat Reader 程序只能正确处理 PDF 文件中 使用 Unicode 编码的书签, 而 CCT 或 (使用 GB/GBK 编码的) CJK 排版文件中通常使用 GBK 编码, 因而生成的 PDF 文件可能会在书签中出现乱码<sup>1</sup>。

用 `dvipdfmx` 程序生成 PDF 文件时, 可以在排版文件中加入命令

```
\AtBeginDvi{\special{pdf:tounicode GBK-EUC-UCS2}}
```

来告诉 `dvipdfmx` 将书签中的汉字转换成 Unicode 编码从而在 AcroBat Reader 中得到正确的书签。而 `pdflatex` 程序则没有提供类似机制, 此时可用 `gbk2uni` 程序对书签文件进行处理。使用 `dvipdfmx` 时也可以用 `gbk2uni` 程序处理书签文件 (这样就不用在排版文件中使用上面的命令)。

`gbk2uni` 程序的使用方法如下:

1. `hyperref` + `pdflatex`:

```
pdflatex myfile.tex
pdflatex myfile.tex
gbk2uni myfile.out
pdflatex myfile.tex
```

2. `hyperref` + `dvipdfmx`:

```
latex myfile.tex
latex myfile.tex
gbk2uni myfile.out
latex myfile.tex
dvipdfmx myfile
```

汉字中含有  $\text{\TeX}$  特殊字符时经 `cctconv` 程序处理后也会在书签文件中产生不正确的汉字 (参看 §3.3), `gbk2uni` 程序也负责对这些字符进行转换。

## §4 Linux 版本

我们为新版 CCT 制作了 Linux RPM 包, 可从 `ftp://ftp.cc.ac.cn/pub/Linux/` 下载。

CCT 的 RPM 包中提供一组脚本文件 `ccttex`, `cctlatex`, `cctamstex`, `cctpdftex` 和 `cctpdflatex`, 它们可分别代替 `tex`, `latex`, `amstex`, `pdftex` 和 `pdflatex` 命令用于 CCT 文件的排版。这些脚本自动完成排版文件的预处理、排版和后处理, 包括以下三个步骤:

---

<sup>1</sup>注: 为正确处理含汉字的书签文件, 必须在 `hyperref` 包中使用 `CJKbookmarks` 选项。

1. 预处理: 如果排版文件的扩展名为“.ctx”则调用 CCT 的预处理程序, 否则调用 `cctspace` 和 `cctconv` 对排版源文件进行预处理。
2. 排版: 调用 `tex`, `latex`, `amstex`, `pdftex` 或 `pdflatex` 进行排版。
3. 后处理: 调用 `gbk2uni` 对书签文件进行后处理。

这些脚本支持可选项“-s”, 它使得预处理时不运行 `cctspace` 程序, 而只运行 `cctconv` 程序。

我们建议在安装 CCT 系统的同时安装 CJK 系统, 以便能够使用新 CCT 系统的所有功能。<ftp://ftp.cc.ac.cn/pub/home/zlb/hack/CJK/> 处提供了一组 CJK 的 RPM 包, 安装 CJK 系统有困难时可以试试它们 (如果系统中已经安装了 CJK 系统及字库, 则在安装这些包前应先备份好重要的配置文件)。这些包都是 `src` 包, 可用 `rpmbuild --rebuild` 命令来生成相应的 `i386` 包。

`CJK-GBKfonts-0.3-3.i386.rpm` 在安装时生成 Type 1 字库, 因此安装前应该事先准备好 TrueType 字库, 并将它们放在一个固定的子目录中。缺省情况下, `CJK-GBKfonts` 将安装四种 Windows 简体中文 TrueType 字体, 它假设 TrueType 字库路径为 `/dos/windows/fonts`, TrueType 字库文件为 `"simsun simhei simkai simfang"` (不含扩展名“.ttf”), 相应的 CJK 字库名为 `"song hei kai fs"` (不含前缀“gbk”)。安装时可以通过设置环境变量 `TTFFONTPATH`, `TTFFONTNAMES` 和 `CJKFONTNAMES` 来改变这些参数, 如:

```
env TTFFONTPATH=/windows/fonts \  
TTFFONTNAMES="simsun simhei simkai simfang simli" \  
CJKFONTNAMES="song hei kai fs ls" \  
rpm -ivh CJK-GBKfonts-0.3-3.i386.rpm
```

注意, `TTFFONTNAMES` 和 `CJKFONTNAMES` 中字库名是一一对应的, 它们的个数一定要相等。此外, 还可以通过环境变量 `GBK FONTS_OPTS` 将命令行参数传递给 `gbkfonts` 程序 (如: `GBK FONTS_OPTS="-f"`)。用 `"rpm -qip CJK-GBKfonts-0.3-3.i386.rpm"` 可以显示一个简要安装说明。

`CJK-GBKfonts` 卸载时会自动删除所创建的文件 (包括 Type 1 字库)。

另外, 目前随  $\text{te}\text{T}_{\text{E}}\text{X}$  发布的 `xdvi` 程序不支持 Type 1 字库, 建议下载并安装 `xdvi` 的最新 CVS 版, 或是用 <ftp://ftp.cc.ac.cn/pub/home/zlb/hack/xdvik/> 中的 `xdvik` 包替换  $\text{te}\text{T}_{\text{E}}\text{X}$  中的 `tetex-xdvi` 包, 以便在预览时可以直接使用 Type 1 字库。