

Unix 的中文列印

沈俊興

元智大學資訊工程學系

statue@netlab.cse.yzu.edu.tw

摘要

在Unix 的環境中，控制印表機輸出最常見的就是ASCII 碼與PostScript。ASCII 碼的輸出不能列印中文，因此當我們要做中文列印，就是要尋求PostScript 圖形列印的途徑，將文字檔轉成PostScript 輸出。中文列印在這幾年有長足的進步，從原本只支援純文字列印的bg5ps、cnprint，慢慢的各種軟體都相繼支援CID-Keyed 字型，如enscript，以做到高品質列印，gs-cjk 的加入讓TrueType 當作CID Font，使得圖形介面的軟體所顯示和列印的效果一致，這讓Mozilla、AbiWord 等軟體有良好的列印品質。幕後排版系統CJK 對中文顯示與列印支援一直不錯。KOffice、OpenOffice 等幕前排版系統則是自行處理的中文列印的問題。本文在討論於FreeBSD 下，各種產生PostScript 檔的方式和對中文的支援，並檢視中文的粗體、斜體、粗斜體等常用需求。也會討論wprint 這種處理有中文問題的PostScript 的應用範圍。

關鍵字: 中文列印

1 簡介

在Unix 的環境中，長久以來已經有許多排版與列印的工具，例如幕後排版系統 $\text{T}_{\text{E}}\text{X}$ / $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ ，列印圖形化語言PostScript，以及常見的PostScript 解譯器GhostScript。然而，這些排版與列印工具在處中文上卻是個難題，主要是中文文字的編碼方式與字型。一般的印表機並不內建中文字型，故當我們要做中文輸出時，直接將檔案送交給印表機列印時，這些印表機不認得中文編碼，也沒有字型可用，就會印出一連串的亂碼。因此將中文字轉成圖形成為目前的解決方案，也分成內嵌中文字型的圖形檔案方式，或是不內嵌中文字型，僅紀錄CID-Keyed 字型名稱，在列印時透過GhostScript 使用CID-Keyed 字型來列印。

PostScript 為Adobe 公司[1] 於1985 年所發表的文件描述技術，這個技術改變了整個印刷工業。PostScript 可以精確的描述繪製平面上任何文字及圖

字體	CID-Keyed 字型
普通	MSung-Light-B5-H
粗體	MSung-Light-Bold-B5-H
斜體	MSung-Light-Italic-B5-H
粗斜體	MSung-Light-BoldItalic-B5-H

表 1: CID Font 對應表

形，現今已經非常普遍的使用在印刷領域。PostScript 是業界文件輸出格式的標準，許多學術論文和排版應刷工作的最終輸出是PostScript，因此也有許多PostScript 印表機，直接可以列印PostScript 的文件而不需要額外的驅動程式。

CID-Keyed 字型[2] 格式的設計是由於大字庫字集的列印需求。CID 是指Character ID，每一個字形外框資料附有一個ID，目前有許多製作CID Font 的方式，最初由Adobe 將教育部所提供的楷書和宋體轉成MOEKai 和MOESung 的CID Font，之後由 gs-cjk [3] 的努力，讓TrueType 字型當作是CID Font；而Keyed 是指一個Mapping 的CMap 檔案，內容包括對應碼位，目前由Adobe 提供這些CMap。以MSung-Light-B5-H 這個CID-Keyed 字型為例，MSung-Light 是Adobe 給繁體明體的預設CID Font 名稱，而B5-H 則是代表Big5 編碼的直書CMap。

而粗體、斜體、粗斜體的需求常見於文書處理軟體、網頁等。目前的做法大多是將原字形位移一點以模擬粗體，原字型傾斜0.167 度來模擬斜體，粗斜體則是粗體加上斜體的效果。相對於CID-Keyed 字型，在CID Font 的結尾加上-Bold、-Italic、-BoldItalic 等以辨識粗體、斜體、粗斜體等效果。如表1: CID Font 對應表。

接下來的章節將討論各種軟體對中文列印的支援方式，然後測試各種軟體所產生的PostScript，最後對目前中文列印做結論。

2 各種軟體的中文列印支援

2.1 GhostScript

GhostScript 是PostScript 印表機模擬程式，透過軟體模擬，可以讓非PostScript 印表機，也能印出PostScript 文件。GhostScript 的工作是轉譯印表指令，將PostScript 指令，轉譯成各種印表機專用的低階列印指令，送印表機執行。因此，GhostScript 需要知道各種印表機的指令語言，這就是所謂的驅動程式。

如果使用CID-Keyed 字型來產生PostScript，粗體、斜體、粗斜體等支援可以直接透過PostScript 的語言來達成，假設我們已經有MSung-Light 這個CID Font，要使用粗體可以在GhostScript 中建立MSung-Light-Bold 這樣

子的資源檔，來將MSung-Light 模擬成粗體。斜體、粗斜體也是同樣的方式。

```
%!PS-Adobe-3.0 Resource-CIDFont
%%BeginResource: CIDFont (MSung-Light-Bold)
/MSung-Light-Bold
/MSung-Light /CIDFont findresource
16 dict begin
  /basecidfont exch def
  /basefont-H /.basefont-H /Identity-H [ basecidfont ] composefont def
  /basefont-V /.basefont-V /Identity-V [ basecidfont ] composefont def
  /CIDFontName dup basecidfont exch get def
  /CIDFontType 1 def
  /CIDSystemInfo dup basecidfont exch get def
  /FontInfo dup basecidfont exch get def
  /FontMatrix [ 1 0 0 1 0 0 ] def
  /FontBBox [
    basecidfont /FontBBox get cvx exec
    4 2 roll basecidfont /FontMatrix get transform
    4 2 roll basecidfont /FontMatrix get transform
  ] def
  /cid 2 string def
  /BuildGlyph {
    gsave
    exch begin
      dup 256 idiv cid exch 0 exch put
      256 mod cid exch 1 exch put
      rootfont /WMode known { rootfont /WMode get 1 eq } { false } ifelse
      { basefont-V } { basefont-H } ifelse setfont
      .03 setlinewidth 1 setlinejoin
      newpath
      0 0 moveto cid false charpath stroke
      0 0 moveto cid show
      currentpoint setcharwidth
    end
    grestore
  } bind def
  currentdict
end
/CIDFont defineresource pop
%%EndResource
%%EOF
```

並建立MSung-Light-Bold-B5-H.gsf 來代表整個CID-Keyed 字型。

```
/MSung-Light-Bold-B5-H
/MSung-Light-Bold (MSung-Light-Bold)
/B5-H (CMap/B5-H)

1 index /CMap resourcestatus
{pop pop pop}
{runlibfile} ifelse
/CMap findresource

3 1 roll
1 index /CIDFont resourcestatus
{pop pop pop}
{runlibfile} ifelse
/CIDFont findresource

[ exch ] composefont pop
```

最後加入將這些.gsf 加入Fontmap.GS。

```
/MSung-Light-Bold-B5-H (MSung-Light-Bold-B5-H.gsf) ;
```

依照上面的方式完成後，就可以得到MSung-Light-B5-H、MSung-Light-Bold-B5-H、MSung-Light-Italic-B5-H、MSung-Light-BoldItalic-B5-H等CID-Keyed 字型，透過GhostScript 來列印就會有粗體、斜體的效果。

2.2 cnprint

cnprint [4] 是用C 語言寫的程式，使用HBF(Hanzi Bitmap Font) 字型將文字檔轉成PostScript 檔案。由於HBF 是點陣字型，本身在放大上會造成鋸齒的問題，將字型轉成PostScript 後，品質也會有相同的問題。cnprint 在目前已經支援部份的TrueType 的功能，但並不理想。

2.3 bg5ps

bg5ps [5] 是一個使用巨蟒描述語言寫的程式，利用TrueType 字型將文字檔轉成PostScript 檔案。bg5ps 的做法是直接將TrueType 字型內嵌到postscript 檔案中，由於已經把中文字的樣式轉成PostScript 的控制語言，所以在列印上就像是列印圖形一般。但是缺點是字型樣式已經內嵌，沒有辦法更換成其他的字型，無法搜尋字串，且當內嵌的字增多的時候，檔案也會急速的增大。

2.4 enscript

enscript [6] 透過CID-Keyed 字型將文字檔轉成PostScript 檔案。當CID-Keyed 字型逐漸成形後，透過CID-Keyed 字型會是不錯的方案。由於CID-Keyed 字型的特性，讓轉出來的PostScript 檔案會比較小，而品質由CID Font 所控制，如果使用TrueType 字型，列印出來的效果跟直接內嵌TrueType 是相同的。字串搜尋的功能也可以使用。

2.5 Mozilla

Mozilla [7] 是一套跨平台免費開放原始碼的瀏覽器，任何人都可以自由下載來使用，而且有非常多先進的設計。Mozilla 提供了設定CID-Keyed 字型的方式來列印文件。不過只能根據單一編碼設定單一的CID-Keyed 字型，並沒有粗體、斜體、粗斜體的效果，也就是說網頁上的粗體、斜體、粗斜體效果在列印到PostScript 檔案時會消失，會變成普通字體，但是中文字不會變成亂碼。

```
pref("print.postscript.nativecode.zh-TW", "big5");  
pref("print.postscript.nativefont.zh-TW", "MSung-Light-B5-H");
```

將HTML 轉成PostScript 還有像html2ps 的指令模式，但是中文的支援上仍然有問題，而且在排版與CSS2 等支援上，Mozilla 仍然有較大的優勢。

2.6 wprint

wprint [8] 是用來過濾 Mozilla、Htmldoc 和Netscape 所產生的PostScript，讓這些PostScript 使用TrueType 以列印各種語言。由於能和 wprint 搭配用來處理的軟體不多，所以目前使用上並不普遍。必須修改wprint.conf 加上字型的設定與輸出方式。

```
fontpath:/usr/local/share/fonts/TrueType/  
default:EUC-TW:msung.ttf=zh_TW.Big5:wprint.ps
```

將 Mozilla 的列印命令改為 wprint 後，當在 Mozilla 中選擇列印到檔案時，就會出現wprint.ps 這個處理過的PostScript 檔案。採用 wprint 就不需要設定CID-Keyed 字型。wprint 直接讀取TrueType 的資料將字形寫入PostScript，既可列印中文，也會保留 Mozilla 的網頁版面。

2.7 AbiWord

AbiWord [9] 的出現提供了另一種列印的方案。AbiWord 的顯示採用Core X11 Fonts 系統[10]，由於有日本人寫的xft 模組[11]，因此顯示介面上可以顯示粗體、斜體、粗斜體，在列印的部分使用Core X11 Fonts 與CID-Keyed 字型的對應，透過簡單的設定，讓-medium-r 對應普通字型，-bold-r 對應-Bold，-medium-i 對應-Italic，-bold-i 對應-BoldItalic，就可以在同一個頁面中顯示並列印粗體、斜體等，達到所見即所得的效果。

```
4  
MSung-Light-B5-H, -default-ming-medium-r-normal--0-0-0-0-c-0-big5-0, 880, 120, 1000  
MSung-Light-Bold-B5-H, -default-ming-bold-r-normal--0-0-0-0-c-0-big5-0, 880, 120, 1000  
MSung-Light-Italic-B5-H, -default-ming-medium-i-normal--0-0-0-0-c-0-big5-0, 880, 120, 1000  
MSung-Light-BoldItalic-B5-H, -default-ming-bold-i-normal--0-0-0-0-c-0-big5-0, 880, 120, 1000
```

2.8 qt3

KOffice [12] 是由KDE 所開發的辦公室套件，包括許多常用的文書處理軟體。KOffice 是由 qt3 函式庫所開發。以fontconfig [13] 作為字型設定的方式，但是fontconfig 對於斜體、粗斜體的支援還不完全，在中文列印上的測試只出現中文普通字。

2.9 CJK

CJK [14] 是由Werner Lemberg 所開發的一個 \LaTeX [15] 套件，專門用來排版中日韓(CJK)文件。CJK 對於中文列印的支持一直都十分良好。CJK 用來處理中文字，讓中文字能正確處理，而 \LaTeX 則是高品質的幕後排版軟體，也可以輸出PostScript。

以 CJK 的設定檔c00msung.fd 為例，以下就是普通字體、粗體、斜體、粗斜體的設定方式。粗體是模擬出來的，CJK 做一點位移達到加粗的變化，與ps 字型本身無關，因此使用普通字體的msung?.tfm 即可。斜體則是由ps 字型本身就能提供傾斜、拉長、加寬的變化，透過tff2tfm 做出msungs?.tfm。粗斜體則是粗體與斜體的綜合效果。

```
\DeclareFontShape{C00}{msung}{m}{n}{<-> CJK * msung}{}  
\DeclareFontShape{C00}{msung}{bx}{n}{<-> CJKb * msung}{\CJKbold}  
\DeclareFontShape{C00}{msung}{m}{it}{<-> CJK * msungs}{}  
\DeclareFontShape{C00}{msung}{bx}{it}{<-> CJKb * msungs}{\CJKbold}
```

在Unix 底下還有cwTeX [16] 與ChiTeX [17] 等中文 \LaTeX 套件，用法與上面相同。

2.10 dvipdfmx

dvipdfmx [18] 主要的功能是轉換 \LaTeX 產生的DVI 檔成為不內嵌中文字型的PDF 檔。PDF 是將難以處理的PostScript 碼精化成一個頁面與物件的資料庫，它包含字型、影像及經壓縮的向量物件。不內嵌中文字型可以使檔案小很多，非常方便網路上的流通，在Acrobat Reader 和xpdf [19] 都可以正常閱覽， \LaTeX 上的高品質排版功能也不會被破壞。

```
msung@Big5@ B5-H :0:!msung.ttf  
msungs@Big5@ B5-H :0:!msung.ttf,Italic  
msung@Big5@ B5-H :0:!msung.ttf,Bold  
msungs@Big5@ B5-H :0:!msung.ttf,BoldItalic  
% 去掉 ! 會嵌入 TTF
```

2.11 OpenOffice

OpenOffice [20] 是昇陽公司[21] 一套重量級的辦公室整合軟體，包含了許許多多常用的辦公室工具。除了介面使用Core X11 Fonts 來顯示外，字型的處理和列印都是本身自行提供的功能。利用 OpenOffice 的列印管理程式安裝好字型後，在顯示與列印上都會自動的處理好，使用者不需要額外的設定。

OpenOffice 也只支援普通字體的列印。

名稱	檔案大小(byte)	頁數	使用環境	使用字型
test.big5	11401			
dvipdfmx	20916	5	X Window	CID-Keyed Font
enscript	38474	4	Console	CID-Keyed Font
Mozilla	211569	5	X Window	CID-Keyed Font
AbiWord			X Window	CID-Keyed Font
cnprint	1729316	5	Console	HBF
CJK	4268989	5	X Window	TrueType
kword	18422813	4	X Window	TrueType
wprint	22896352	5	X Window	TrueType
OpenOffice	25716441	4	X Window	TrueType
bg5ps	35576903	6	Console	TrueType

表 2: 各軟體比較表

3 測試結果

以一個包含5607 不同中文字的test.big5 純文字檔作為測試的依據，並以GPL 的文鼎PL上海宋[22] 作為測試的TrueType，與CID Font。測試結果如表2: 各軟體比較表。

```
cat test.big5 | bg5ps > bg5ps.ps
enscript -X big5 -f MSung-Light-B5-H@l2 -o enscript.ps test.big5
cnprint -big5 -f=ntu_kai48.hbf test.big5 -w -o=print.ps
```

Mozilla 透過網頁瀏覽來列印PostScript。kword 讀入純文字檔後來列印成PostScript。CJK 加上簡單的tex 格式後，用latex 產生.dvi，然後用dvips 產生PostScript。dvipdfmx 將 CJK 產生的.dvi 做成不內嵌中文字型的PostScript。KWord 這套軟體就是KOffice 裡的文書編輯軟體。AbiWord 在測試時由於無法讀入fonts.dir，所以沒有做出PostScript。

由於測試檔中包含特殊的Big-5 日文，在測試 CJK 時常用的ttf2tfm 或ttf2pk 並沒有對應表，因此使用ttf2pt1 [23] 來產生所需要的tfm 檔案。這些特殊的Big-5 日文在大部份的軟體中並沒有辦法顯示，列印出來的結果也有問題。

4 結論

部份軟體直接使用TrueType 來列印，直接使用的好處就是可以不依賴其他軟體提供的介面來使用TrueType，有完全的自主權，大型的套裝軟體都採用如此的作法。原本以為CID-Keyed 字型的出現會改變中文列印的世界，因為使用CID-Keyed 字型最大的好處就是檔案小，適合出版，在網路上流傳等，真正使用CID-Keyed 字型的只有部分軟體。

但是粗體、斜體、粗斜體卻沒有普遍的被重視，只有 AbiWord 和 CJK 的支援最好。fontconfig 在粗體、斜體、粗斜體的部分還不支援，所以 KDE3 和 gnome2 相關的軟體應該都還不支援這方面的顯示和列印，不過目前已經有不少非官方的修補檔，還需要多花點時間去注意和測試，因為 fontconfig 有取代目前廣泛使用 Core X11 Fonts 的趨勢。

因此，在純文字模式底下要列印中文，可以採用 cnprint、bg5ps、enscript 等方式。在 HTML 列印使用 Mozilla 可以採用設定 CID-Keyed 字型，或是使用 wprint 過濾器。辦公室套裝軟體的部份，對中文都有一定的支援，在中文列印上算是堪用了。

參考文獻

- [1] Adobe Solutions Technical Resources
<http://partners.adobe.com/asn/techresources.jsp>
- [2] CID-Keyed Fonts
<http://partners.adobe.com/asn/tech/type/cidfonts.jsp>
- [3] gs-cjk
<http://www.gyve.org/gs-cjk/>
- [4] CAI's Software Page - cnprint
<http://www.neurophys.wisc.edu/cai/software/>
- [5] bg5ps
<http://students.washington.edu/cschin/>
- [6] enscript
<http://www.gnu.org/software/enscript/enscript.html>
- [7] Mozilla.org
<http://www.mozilla.org/>
- [8] World Print
<http://ttt.esperanto.org.uy/programoj/angle/wprint.html>
- [9] AbiWord - Word Processing for Everyone
<http://www.abisource.com/>
- [10] Fonts in XFree86
<http://www.xfree86.org/dawes/4.3.0/fonts.html>
- [11] After X-TT Project
<http://x-tt.sourceforge.jp/>

- [12] The KOffice Project
<http://koffice.kde.org/>
- [13] Fontconfig.Org
<http://www.fontconfig.org/>
- [14] The CJK package for L^AT_EX
<http://cjk.ffii.org/>
- [15] The teTeX Homepage
<http://www.tug.org/teTeX/>
- [16] cwTeX
<http://ceiba.cc.ntu.edu.tw/tmwu/>
- [17] ChiTeX
<http://www.math.ncu.edu.tw/yih/>
- [18] The dvipdfmx Project
<http://project.ktug.or.kr/dvipdfmx/>
- [19] Xpdf
<http://www.foolabs.com/xpdf/>
- [20] OpenOffice .org P^SPrint - a Unix print solution
<http://gsl.openoffice.org/psprint/index.html?JServSessionIdservlets=djfkxhbps1>
- [21] Sun Microsystems
<http://tw.sun.com/>
- [22] 文鼎科技
<http://www.arphic.com.tw/>
- [23] ttf2pt1 Project
<http://ttf2pt1.sourceforge.net/>