Avoiding Type 3 Bitmap Fonts in PDFs Created From DVI Files

Rick Holbert

Mar 10, 2002

1 Background Information

When people talk about PostScript, it might refer to one of the following several things:

- A Page Description Language

  The PostScript specification was created by Adobe in 1985, which is open. Though Adobe owns the copyright to the list of the operators and the written specification of the language, the general idea of such a page description language has fallen into the public domain. That's why you see so many vendors supporting it in the market, for example, our community has GhostScript and Ghostview (or newly GV) which support this open specification.

- An Interpreter

  The PostScript interpreter can translate PostScript instructions for devices such as a display on the computer screen, printers, etc. GhostScript is such an interpreter which is free under GPL. GNUStep adopted a rending engine based on PostScript technology which shows all the screen graphics as a desktop GUI environment.

  The interpreter can also be embedded into certain types of hardware, like a printer (or other devices used for printing purposes). Such a printer is called a PostScript printer.

- A Programming Language

  According to the specification, all the source code of PostScript programs should be written in ASCII, thus, PostScript programs can be read by and modified by a programmer. This also means that PostScript programs are portable among various computer systems.

- A Registered Trademark ^{TM} of Adobe Systems Inc.

  As Adobe owns the copyright and the trademark of PostScript, any other interpreters that can execute the PostScript programming language can not use the trademarked name PostScript. That's why the GPL programs GhostScript and Ghostview were so named.

PostScript is vector-based, this means — that unlike bitmap graphics — the size of a PostScript file is only related to the complexity of the picture per se, and not related the physical size of the picture. This idea was also adopted by the new W3C scalable vector graphics standard (SVG).

Technically, PostScript is a stack-based programming language, it has strange grammar rules, and strange structure consists of *prolog* and *script*. But once you get used to them, then Programming in PostScript becomes more enjoyable.

In order to increase its interpretation speed, PostScript defined a rich set of objects, including fonts. This enables fast rendering speed, as the pre-defined fonts do not need to be reinterpreted from scratch each time, using the glyph's source code. As the history of PostScript is quite long, various font types are available. For more details of the font types, please take a look at Page 322 of *PostScript Language Reference*, 3rd edition http://partners.adobe.com/asn/developer/technotes.

2 Problem

One of my pet peeves is the unnecessary use of Type 3 bitmap fonts in Portable Document Format (PDF) files created from DVI files. The resulting PDF files are large, ugly, slow to display and print.

When I saw that the first issue of FREE SOFTWARE Magazine was using Type 3 fonts in its PDF files, I emailed the editor, Hong Feng, and volunteered to convert all the articles into quality PDFs using Type 1 outline fonts.

3 Discussion

The advantages of Type 1 fonts are three fold. First, they look a lot nicer on the screen. Second, they often result in smaller PDF files. Third, they display, and print faster.

Now that we know about the disadvantages of Type 3 bitmap fonts, and the advantages of Type 1 outline fonts, lets look at some solutions.

4 The Easy Solution

Hong Feng sent me a single DVI file for conversion testing. I successfully converted it into a nice looking PDF file, and mailed it back to him. Then he tarred and gzipped all the DVI files for Issue 2, and emailed them to me. I gunzipped, and untarred the DVI files, and issued the following commands to convert each of them into PDF files:

```
$ dvips -Pwww file_name.dvi -o file_name.ps
$ ps2pdf file_name.ps
```

That's it! The first command tells dvips to convert the DVI file into a PostScript file, and the second line tells GhostScript to distill the PostScript file into a PDF file.

The key here is the `-Pwww` (printer driver) flag, which tells dvips to try to use Type 1 outline fonts in lieu of Type 3 bitmap fonts.

This technique works well with many of the common fonts, including the Computer Modern (CM) fonts. Also, some versions of dvips can use `-Ppdf`, but `-Pwww` is more widely supported.

This is fine, but what about some of the other more common fonts, such as the European Common (EC) fonts?

5 European Challenge

The last challenge I faced was the Table of Contents file. Hong Feng had used European Common fonts when he encoded the Table of Contents DVI because some of the authors were European, and had special characters in their names. My `-Pwww` trick didn't help here, because my default install of tetex didn't include any European Common Type 1 fonts.

6 More Complicated Solution

I remembered reading about the textrace project on Freashmeat (http://www.freshmeat.net). Textrace was designed to convert TEX METAFONTS into Type 1, `.pfb` outline fonts.

While visiting the textrace web site (http://www.inf.bme.hu/~pts/textrace), I discovered the cm-super font package. This package, developed by Vladimir Volovich, contains Type 1 versions of the entire EC/TC, ECC and LH fonts. All European and Cyrillic characters are supported. The following standard LATEX encodings are included: T1, TS1, T2A, T2B, T2C, and X2.

So, armed with the LATEX source for the Table of Contents, and this new information, I downloaded `cm-super.tar.gz` from the packages directory at

http://ftp.yars.free.net/pub/software/tex, and installed it by following the instructions found in the REAME file.

I added `\usepackage{pslatex}` to the preamble, ran toc.tex through LaTeX, and used `dvi -Pwww toc.dvi -o toc.ps` to produce a PostScript file. I then used `ps2pdf toc.ps` to produce a PDF file, and it worked!

7 Summary

We no longer have to live with ugly PDF documents. Creating PDF files using Type 1 fonts results in documents that are easy on the eyes, smaller in size, faster to print and display.

8 About the Author

Rick Holbert is a Unix System Administrator, currently working for the Office of Information Technology at The Ohio State University and is also a member of the Central Ohio Linux Users Group and the Columbus Computer Society.

Rick is a 1982 graduate of the United States Air Force Academy and an Ohio native.

He has several years of computing experience. His first computer was an Apple][+ that he modified so he could run CP/M. He started using Unix in 1985, and installed his first copy of GNU/Linux in 1995.

Some of his passions include amateur radio, and micro publishing. He presented some of his ideas at the 2001 O'Reilly Open Source Convention. He hopes to present again at OSCON 2002.

He's published several free electronic books at http://rholbert.colug.net