HTMLDOC 1.8 User's Guide

Michael R. Sweet

Table of Contents

Table of Contents

Table of Contents

Introduction

About This Software

This document describes how to use the *HTMLDOC* software, version 1.8. *HTMLDOC* is a HTML document processing program that generates indexed HTML, Adobe® PostScript™, and Adobe Portable Document Format (PDF 1.2) files suitable for printing or online viewing.

No restrictions are placed upon the output produced by *HTMLDOC*.

History

Like many programs *HTMLDOC* was developed in response to a need my company had for generating high−quality documentation in printed and electronic forms. For a while we used FrameMaker® and a package from Silicon Graphics that generated "compiled" SGML files that could be used by the Electronic Book Technologies documentation products (EBT is now owned by INSO.) When SGI stopped supporting these tools we were stuck. INSO laughed in our faces when we said we could afford up to $10,000 US for an updated book generation product that would work with the IRIX, Solaris, and HP−UX documentation systems. Other solutions were nearly as expensive, or simply not capable of doing the things we needed.

Because of these things I decided to write my own program to generate our documentation. HTML seemed to be the source format of choice since WYSIWYG HTML editors are widely (and freely) available and at worst you can use a plain text editor. We needed HTML output for documentation on our web server, PDF for customers to read and/or print from their computers, and PostScript for our own printing needs.

The result of my efforts is the *HTMLDOC* software which is now available for UNIX® and Microsoft® Windows®. Among other things, this user's guide was produced by *HTMLDOC*.

Why Just HTML?

Some people have asked why this program only deals with HTML input files and is not able to read any Standard Generalized Markup Language (SGML) file. The reasons are numerous but basically boil down to:

1. SGML is a moving target since all SGML documents use Document Type Definition (DTD) files to define what markups are actually supported. Formatting and processing can become a nightmare if the DTD file contains even a single typographic error. Also, this would make the front−end parsing code that generates the document markup tree considerably larger than it already is, not to mention complicating the output code.
2. Tools for SGML file generation cost considerably more than software that generates HTML files. Also, the number of HTML tools is at least an order of magnitude greater than the number of SGML tools!

In the future I may add support for XML files, but at present XML has very little in the way of tools. Time will tell...

Organization of This Manual

This manual is organized into the following chapters and appendices:

If you have downloaded one of the many precompiled binaries from our FTP server then you can skip chapter 1.

If You Have Problems

If you have difficulty using or compiling *HTMLDOC* please send EMail to "mike@easysw.com". Additional information can also be found at the *HTMLDOC* web page at "http://www.easysw.com/~mike/htmldoc".

Copyright and Trademark Information

Incorporated.

Microsoft, Windows, Windows 95, and Windows NT are registered trademarks of Microsoft Corporation.

SPARC is a registered trademark of SPARC International, Inc.

Solaris is a trademark of Sun Microsystems, Inc.

IRIX is a trademark of Silicon Graphics, Inc.

Digital is a trademark of Digital Equipment Corporation.

UNIX is a registered trademark of the X/Open Company, Ltd.

This software is based in part on the work of the Independent JPEG Group.

1 − Compiling HTMLDOC

This chapter describes the steps needed to compile *HTMLDOC* on your system.

Getting a Precompiled Executable

If you don't think you're up to compiling *HTMLDOC*, or you don't have the compiler and libraries listed below, consider downloading a precompiled version of *HTMLDOC*. Precompiled binaries are currently available for the following systems:

- Digital™ UNIX® 4.0 or higher
- HP−UX 10.20 or higher
- IRIX™ 5.2 or higher
- Linux x86 2.0 or higher
- Solaris® SPARC® 2.5 or higher
- Solaris x86 2.5 or higher
- Windows 95® or Windows NT® 4.0 or higher

Requirements

HTMLDOC requires the following software and libraries:

- C++ compiler (GCC will work)
- JPEG library, version 6 or higher
- PNG library, version 1.0.1 or higher

- GZIP library (zlib), version 1.0.4 or higher
- Fast Light Tool Kit (FLTK), version 19990107 or newer
- X11 libraries, R5 or higher (UNIX only).

The GZIP library is used for reading PNG image files as well as writing compressed PDF files. GIF reading support is provided by *HTMLDOC* source code.

The JPEG library is used for reading JPEG image files as well as writing JPEG−compressed images in Level 2 PostScript and PDF output.

For the Microsoft Windows version of *HTMLDOC* you'll probably need Microsoft Visual C++ 5.0 or higher (other PC compilers *may* work; I didn't have much luck with Borland C++ 5.02).

Compiling under UNIX

If you are compiling for Windows, see "Compiling with the Visual C++ Project File".

HTMLDOC is built from a single Makefile in the distribution's main directory (*htmldoc−1.8*). To configure the Makefile for your system you must run the *configure* script:

```
% ./configure Enter
```

To compile *HTMLDOC* simply run the "make" command in the *HTMLDOC* directory. If you get any fatal errors please send a copy of the make/compiler output to "mike@easysw.com" for assistance. Please note the version of *HTMLDOC* that you are using as well as any pertinent system information (operating system, OS version, compiler, etc.)

Installing the Software

The Makefile built by the *configure* script supports installation of the program and man pages under */usr/local* or a directory provided to the *configure* script using the "−−prefix=directory" option.

To install *HTMLDOC* simply run the "make install" command as root.

Compiling with the Visual C++ Project File

A sample project file for Visual C++ 5.0 is included in the source distribution in the file *htmldoc.dsp*. You will need to change the include directories and libraries to point to the directories containing the JPEG, PNG, ZLIB, and FLTK libraries.

2 − Guidelines for HTML Source Files

HTMLDOC is capable of processing most HTML files. This chapter discusses the requirements for HTML files to be correctly processed by *HTMLDOC*

General Requirements

Since *HTMLDOC* is designed as a documentation generation program, it expects to have chapters and headers.

Note: If you are converting a generic web page you must select the "−−webpage" option on the command−line or choose "Web Page" as the document type in the GUI.

Page Breaks

To force a page break use the HR markup with the option "BREAK":

```
<HR BREAK>
```

Chapters

All chapters start with a top−level heading (H1) markup. Any headings within a chapter must be of a lower level (H2 to H7). Each chapter starts a new page (the next odd−numbered page if duplexing is selected.)

Headings

The headings you use within a chapter must start at level 2 (H2). If you skip levels the heading will be shown under the last level that was known. For example, if you use the following hierarchy of headings:

```
<H1>Chapter Heading</H1>
...
<H2>Section Heading 1</H2>
...
<H2>Section Heading 2</H2>
...
<H3>Sub−Section Heading 1</H3>
...
<H4>Sub−Sub−Section Heading 1</H4>
...
<H4>Sub−Sub−Section Heading 2</H4>
...
<H3>Sub−Section Heading 2</H3>
...
<H2>Section Heading 3</H2>
...
<H4>Sub−Sub−Section Heading 3</H4>
...
```

the table−of−contents that is generated will show:

Chapter Heading

- ♦ Section Heading 1
- ♦ Section Heading 2
  - ◊ Sub−Section Heading 1
    - · Sub−Sub−Section Heading 1
    - · Sub−Sub−Section Heading 2
  - ◊ Sub−Section Heading 2
    - · Sub−Sub−Section Heading 3
- ♦ Section Heading 3

Unsupported or Restricted HTML Features

The following HTML features are either not supported or have limited support in this release of *HTMLDOC*.

Embedded Objects

Only embedded HTML files are supported using the EMBED tag.

Fonts

Limited typeface specification is currently supported. The "Arial" typeface is mapped to "Helvetica" to ensure portability across platforms and for older PostScript printers. All other unrecognized typefaces are silently ignored.

Forms

Forms are not yet supported when generating PostScript and PDF files.

Frames

HTMLDOC does not support frames.

Image Maps

Image maps are not exported to HTML or PDF files.

Links

External URL links are fully supported for HTML and PDF output. Internal links are supported in HTML and PDF output.

When generating PDF files, local file links will be converted to external file links for the PDF viewer instead of URL links. That is, you can directly link to another local PDF file from your HTML document.

Scripts and Applets

All scripts and applets are silently stripped from the output.

Style Sheets

Style sheets are not yet supported.

Tables

Currently only the HTML 3.2 varient of tables is supported. The CAPTION, THEAD, TFOOT, and TBODY tags are silently ignored, and vertical alignment and spanning (ROWSPAN and VALIGN) are not supported.

3 − Generating Documents from the GUI

This chapter describes how to generate document files from a given set of HTML "source" files using the *HTMLDOC* GUI.

Starting the HTMLDOC GUI

To start the *HTMLDOC* GUI under UNIX, type:

```
# htmldoc Enter
```

To start the *HTMLDOC* GUI under Windows, choose *HTMLDOC* from the *Start* menu (Figure 1.)

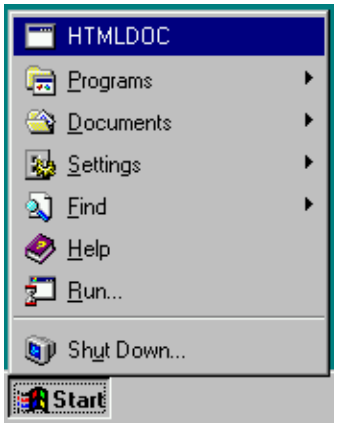

Figure 1 − Starting HTMLDOC under Windows

The HTMLDOC GUI

The *HTMLDOC* GUI (Figures 2 through 7) is contained in a single window showing the input, output, and generation options. At the bottom are buttons to load, save, and generate documents.

Document File Operations

HTMLDOC stores the HTML files, settings, and options in "book" files. The buttons on the bottom of the *HTMLDOC* window allow you manage these files and generate formatted documents.

Starting a New Document

To start a new document, click on the *New* button.

Opening an Existing Document

To open a document you've saved previously, click on the *Open...* button.

Saving the Current Document

To save the current document, click on the *Save* button. If you have never saved the document before or would like to save it with a new filename, click on the *Save As...* button instead.

Note: Saving a document is not the same as *generating* a document. The book files saved to disk by the *Save* and *Save As...* buttons are *not* the final HTML, PDF, or PostScript output files. You generate those files by clicking on the *Generate* button described below.

Generating Your Document

To generate your document, click on the *Generate* button. The progress meter at the bottom of the window will show the progress as each page is formatted and written.

Note: Generating a document is not the same as *saving* a document. To save the current HTML files and settings in the *HTMLDOC* GUI, click on the *Save* or *Save As...* buttons instead.

Exiting from the HTMLDOC GUI

To exit from the *HTMLDOC* GUI, click on the *Close* button.

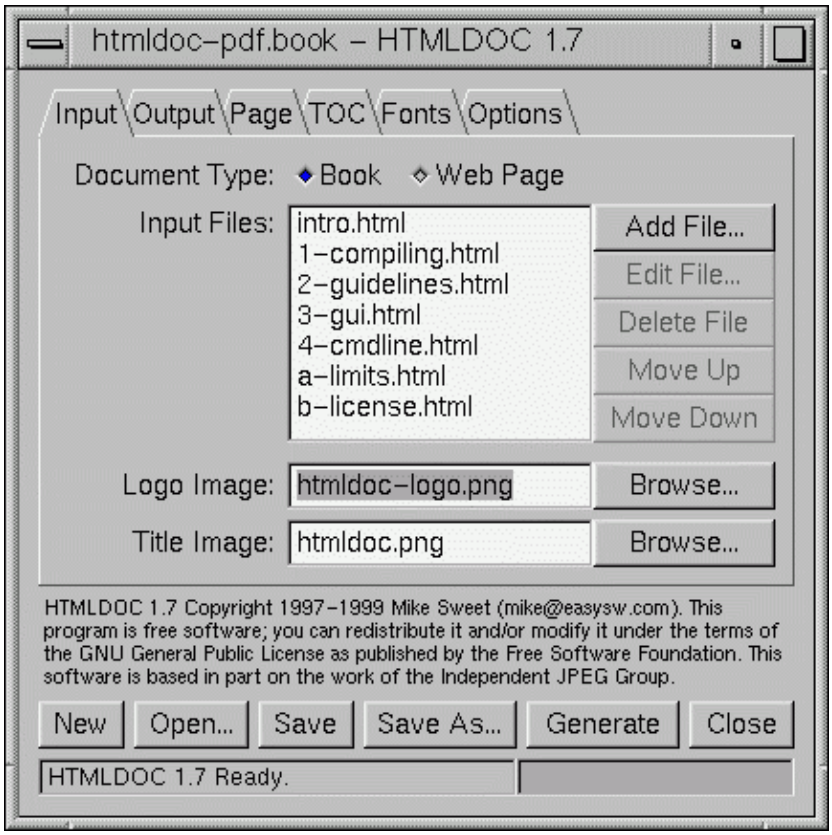

Figure 2 − HTMLDOC Window and Input Tab

The Input Tab

The input tab lists all of the HTML source files that are used to generate the document. You also specify the type of document (book or web page) and the title and logo images in this tab.

Setting the Document Type

Normally *HTMLDOC* generates indexed documents from your structured HTML files. To convert a single unstructured document (or "web page") click on the *Web Page* radio button.

Adding HTML Input Files

Click on the *Add...* button to add an HTML file to your document.

Deleting HTML Input Files

To remove one or more HTML files from your document, click on the file (or drag for multiple files) in the input file list and then click on the *Delete* button. The files are removed from your document but are *not* deleted from your disk.

Editing HTML Input Files

To edit one or more HTML files in your document, click on the file (or drag for multiple files) in the input file list and then click on the *Edit...* button. By default this starts the *nedit* editor under UNIX and the *Notepad* editor under Windows. See "The Options Tab" later in this chapter for details on how to change the editor that is used.

Moving HTML Input Files

To change the order of the input files, click on a file to move (or drag multiple files) in the input file list and then click on the *Move Up* or *Move Down* button.

Selecting a Logo Image

The logo image is shown on the title page of PostScript and PDF output files and in the navigation bar of HTML files. To select a logo image file, click on the *Browse* button. After the standard file selection dialog appears, double−click on the desired image file.

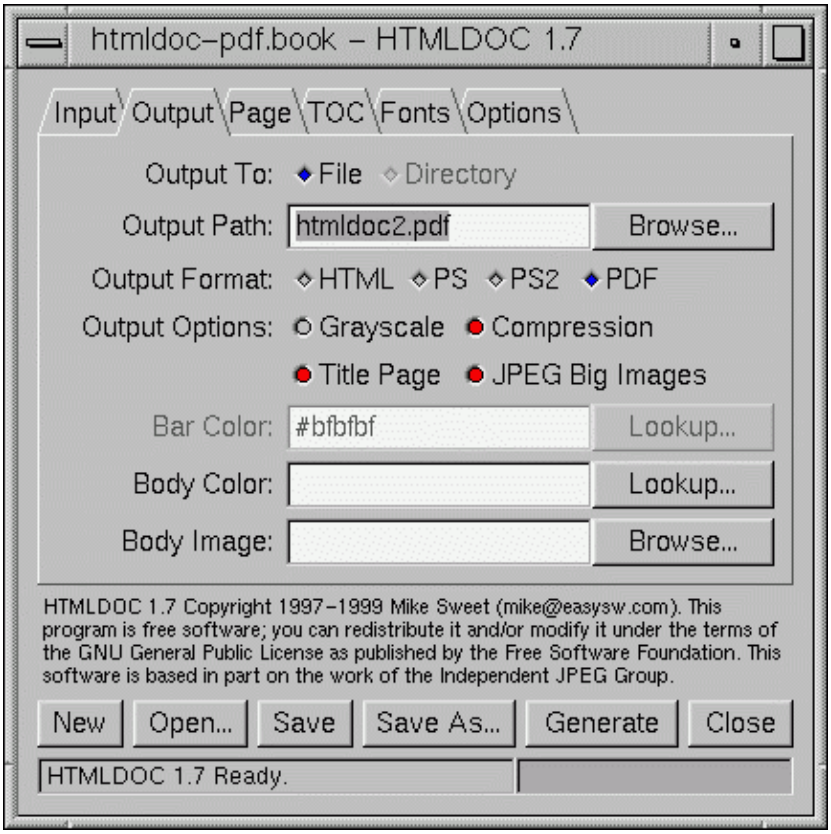

Figure 3 − HTMLDOC Output Tab

The Output Tab

The output tab specifies where your document will be generated, the output format, and some of the output options that apply to all formats.

Selecting File or Directory Generation

HTMLDOC can generate a single HTML or PostScript file or a series of files, one per chapter plus the table of contents (index) file. To select single file output click on the *File* radio button. To generate multiple files to a directory click on the *Directory* radio button.

Selecting an Output File or Directory

The output file is the HTML, PostScript, or PDF file you wish to generate from your HTML files. To select an output file, click on the *Browse* button. After the file selection dialog appears, type the name of the file you would like to create.

Selecting the Output Format

To select an output format, click on the corresponding *Output Type* button. Be careful when generating Level 2 PostScript output, as Level 1 PostScript printers do not support the Level 2 image commands generated by *HTMLDOC* (most printers manufactured in the last 4 years are Level 2).

Note: Choose Level 2 PostScript output for Level 3 PostScript printers.

Selecting Grayscale Output

When generating PostScript or PDF files you can choose to convert all images to grayscale. This is necessary for many Level 1 printers that do not support color images and can reduce the size of output files considerably.

To select grayscale output, click on the *Grayscale* button.

Selecting Compressed Output

PDF files are compressed using Flate (a.k.a. ZIP) compression by default. If you need to view the PDF files produced by *HTMLDOC* with a PDF viewer that does not support this compression, click on the *Compression* toggle button to turn compression off. You may also want to change the PDF Version option.

Disabling the Title Page

A title page is generated for your document by default. To turn the title page off, click on the *Title Page* toggle button.

Using JPEG Compression

HTMLDOC supports JPEG compression of large images when generating Level 2 PostScript and PDF files. To enable JPEG compression, click on the *JPEG Big Images* toggle button. The output quality can be controlled by dragging the *JPEG Quality* slider in the options tab.

Once you have enabled JPEG compression, any color image that cannot be converted to an 8−bit (or less) colormapped image will be JPEG'd. Similarly, any grayscale image that cannot be represented by 16 (or less) shades will be JPEG'd.

JPEG compression can dramatically reduce the size of output files, however with low quality settings the images can look blotchy.

Changing the Navigation Bar Color

To change the color of the navigation bar used in HTML output, type in a color name in the *Bar Color* field or click on the *Lookup...* button to graphically pick a color.

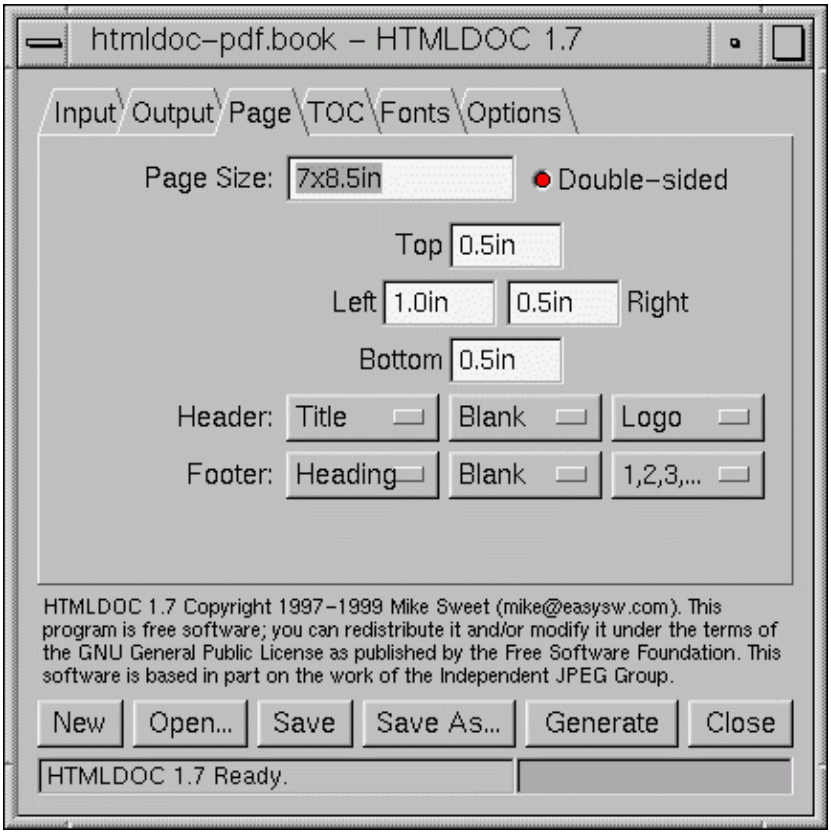

Figure 4 − HTMLDOC Window and Page Tab

The Page Tab

The page tab defines the page header, footer, size, and margins for PostScript and PDF output.

Selecting a Page Size

The page size option is only available for PostScript and PDF output. *HTMLDOC* supports the following standard page size names:

- Letter − 8.5x11in (216x279mm)
- A4 − 8.27x11.69in (210x297mm)
- Universal − 8.27x11in (210x279mm)

To select a custom page size, double−click on the page size text and enter the page width and length separated by the letter "x". Append the letters "in" for inches, "mm" for millimeters, or "cm" for centimeters.

Note: This option does not set the page size on the printer. It only generates pages using the specified size. You must still select the correct page size from your application or printer options.

Selecting Double−Sided Output

To select double−sided (duplexed) output click on the *Double−Sided* toggle button.

Note: This option does not select duplexing on the printer. It only generates pages with the left/right margins swapped on even numbered pages and forces all chapters (and the table−of−contents) to start on an odd−numbered page. You must still select duplexing from your application or printer options.

Setting the Page Margins

The left, right, top, and bottom margins can be changed by clicking in the appropriate text field and entering a new margin. Append the letters "in" for inches, "mm" for millimeters, or "cm" for centimeters.

Customizing the Header and Footer

To customize the header and footer for the document/body pages, select the desired text from each of the option buttons. The leftmost option buttons set the text that is left−justified, while the middle buttons set the text that is centered and the right buttons set the text that is right−justified.

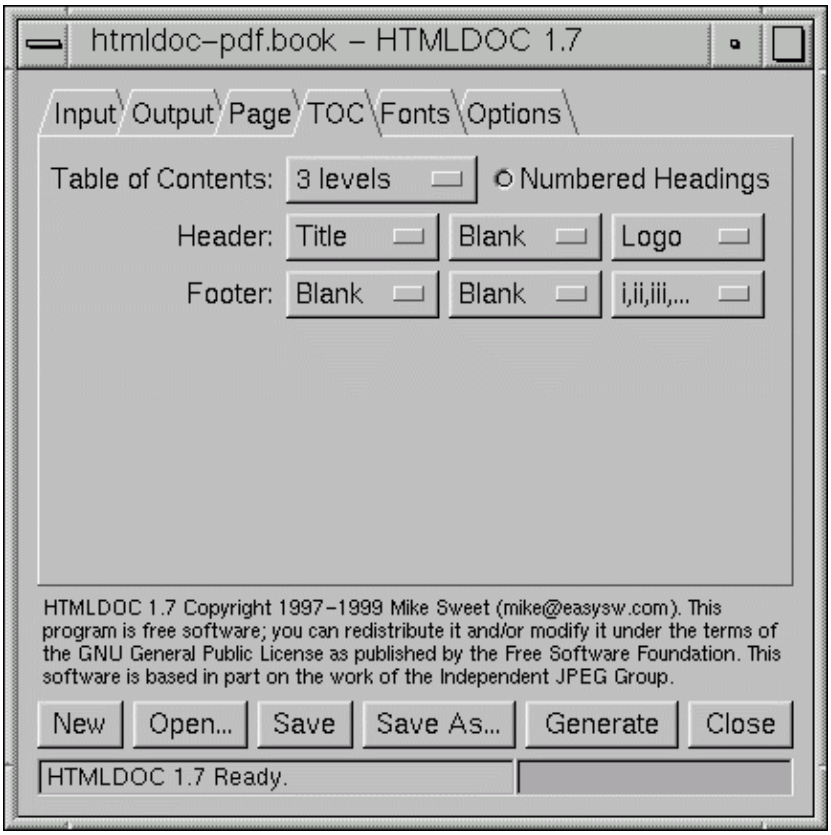

Figure 5 − HTMLDOC Window and TOC Tab

The Table−Of−Contents Tab

The table−of−contents tab defines the number of levels in the table−of−contents and the page header and footer that are used when generating PostScript and PDF files.

Customizing the Table of Contents

To change the number of header levels listed in the table of contents, or to turn off table−of−contents generation entirely, click on *Table of Contents* chooser and select the number of levels desired.

Numbering Table of Contents Headings

To number the headings in your document, click on the *Numbered Headings* toggle button.

Customizing the Header and Footer

To customize the header and footer for the table−of−contents pages, select the desired text from each of the option buttons. The leftmost option buttons set the text that is left−justified, while the middle buttons set the text that is centered and the right buttons set the text that is right−justified.

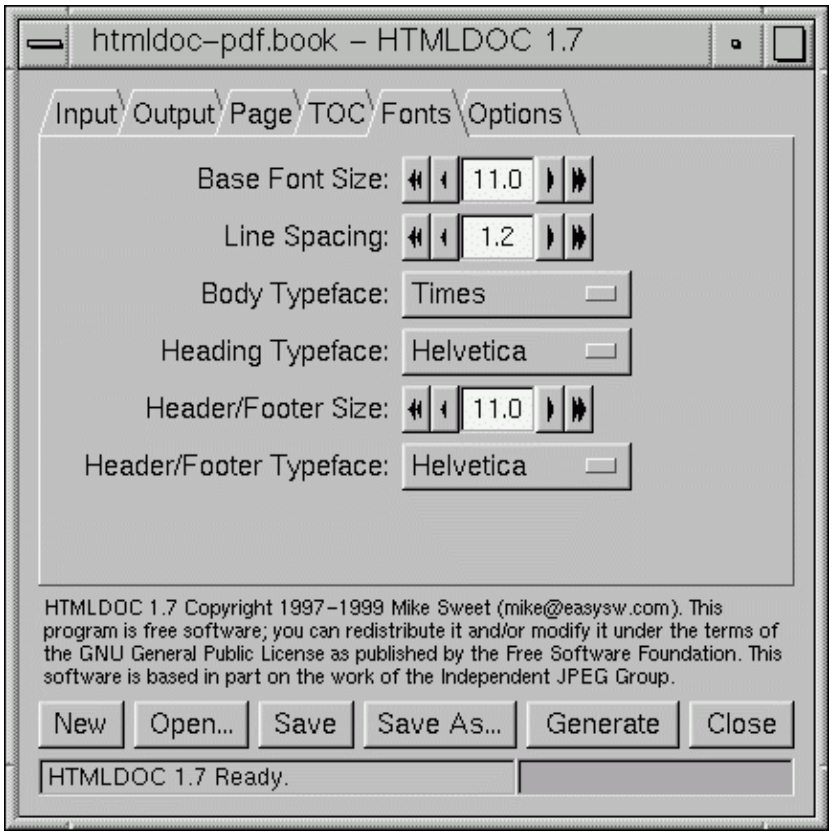

Figure 6 − HTMLDOC Fonts Tab

The Fonts Tab

The fonts tab contains all of the document font options. The default options roughly correspond to those used by most browsers.

Changing the Base Font Size

To change the base font size, click on the left arrow buttons to decrease the font size and the right arrow buttons to increase the font size. The font size value is in points (there are 72 points per inch).

Changing the Line Spacing

To change the line spacing, click on the left arrow buttons to decrease the line spacing and the right arrow buttons to increase the line spacing.

Changing the Body Typeface

The body typeface is the font used for paragraphs and most other text in a document. To change the body typeface click on the chooser and pick the desired typeface.

Changing the Heading Typeface

The heading typeface is the font used for headings. To change the headings typeface click on the chooser and pick the desired typeface.

Changing the Header/Footer Size

To change the header and footer font size, click on the left arrow buttons to decrease the font size and the right arrow buttons to increase the font size. The font size value is in points (there are 72 points per inch).

Changing the Header/Footer Typeface

The header/footer typeface is the font used for headers at the top of the page and footers at the bottom of the page. To change the header/footer typeface click on the chooser and pick the desired typeface.

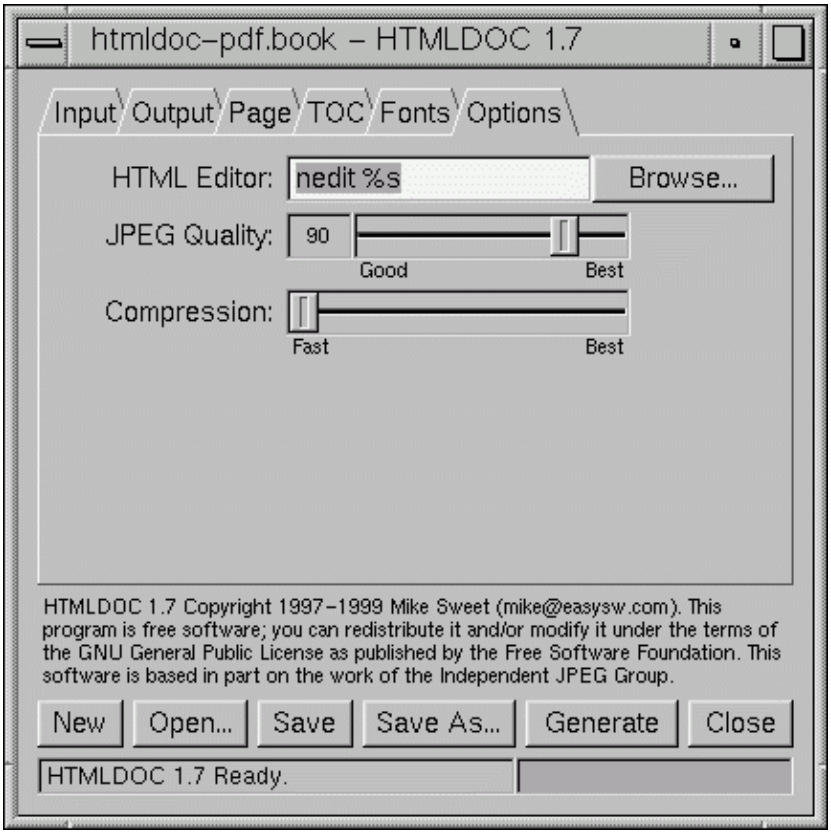

Figure 7 − HTMLDOC Options Tab

The Options Tab

The options tab contains options specific to PostScript and PDF output as well as the current HTML editor.

Changing the HTML Editor Command

To change the HTML editor that is used, type in the program name in the *HTML Editor* field or click on the *Browse...* button. The "%s" is required and is replaced by the file to edit.

NOTE: To use Netscape Communicator as your HTML editor you need to add the "−edit" option before the "%s".

Changing the JPEG Quality

To change the JPEG quality setting, move the mouse pointer over the slider knob and drag the slider using the left mouse button. Release the mouse button when the desired quality is shown.

Changing the Compression Setting

To change the compression setting, move the mouse pointer over the slider knob and drag the slider using the left mouse button. Release the mouse button when the desired level is shown.

Changing the PDF Version

Older versions of Adobe Acrobat Reader cannot handle the Flate compression and named links present in the normal output of HTMLDOC. To change the PDF version, click on the radio button that corresponds to the version of Acrobat you or your users will be using. If you have Acrobat Reader 4.0, choose version 1.2.

4 − Generating Documents from the Command−Line

This chapter describes how to generate one or more document files from a given set of HTML "source" files using the *HTMLDOC* software from the command−line. If you are converting web pages from HTML to PostScript or PDF format, be sure to look at the Converting Web Pages section.

Generating a Single File

To generate a single file containing the entire document, type the following:

```
% htmldoc −f outfile.html infile1.html infile2.html ...
% htmldoc −f outfile.pdf infile1.html infile2.html ...
% htmldoc −f outfile.ps infile1.html infile2.html ...
```

The "outfile.html", "outfile.pdf", and "outfile.ps" arguments are the desired output file. "infile1.html", "infile2.html", etc. are your HTML source files.

By default *HTMLDOC* looks at the extension of the output file to determine the output format. Files ending in ".ps" select Level 2 Adobe® PostScript™ output. For Level 1 PostScript see "Forcing Level 1 Output" below.

Generating Multiple Files

To generate multiple files for the document, type the following:

```
% htmldoc -d outdir -t html infile1.html infile2.html ...
% htmldoc -d outdir -t ps1 infile1.html infile2.html ...
% htmldoc -d outdir -t ps2 infile1.html infile2.html ...
```

The "outdir" argument is the desired output directory. The "−t html", "−t ps1", and "−t ps2" arguments select HTML and PostScript output, respectively. "infile1.html", "infile2.html", etc. are your HTML source files. A separate HTML or PostScript file (doc####.html or doc####.ps) will be created for each chapter (H1 heading) in the document as well as a table of contents file (index.html or index.ps). For HTML output, all local image files that are referenced in the document will be copied to the output directory as well.

Multiple output files are currently not supported for PDF output.

General Options

The following options apply to all output formats.

Numbering the Headings

Some types of documents require paragraph/heading numbers. To enable automatic heading numbering use the "−−numbered" option:

```
% htmldoc --numbered -f outfile.html ...
```

Adding a Logo Image

The logo image is optionally displayed in the page heading of PostScript and PDF output and at the top of the navigation bar along the left side of the page of HTML output. To include a document "logo" use the "−−logo" option to *HTMLDOC*:

```
% htmldoc --logo logo.gif ...
```

The logo file can be of any supported image file type (GIF, JPEG, PNG).

Adding a Title Image

The title image is displayed on the title page. To include a title page image use the "−−title" option to *HTMLDOC*:

```
% htmldoc --title title.gif ...
```

The logo file can be of any supported image file type (GIF, JPEG, PNG).

Converting Web Pages

To convert unstructured HTML documents such as web pages, use the "−−webpage" option to *HTMLDOC*:

```
% htmldoc --webpage ...
```

This is equivalent to using the "−−no−title" and "−−no−toc" options.

Setting the Table of Contents Depth

To set the number of heading levels to show in the table−of−contents use the "−−toclevels" option to *HTMLDOC*:

```
% htmldoc --toclevels # ...
```

The default depth is three levels (H1 to H3). To turn the table of contents off, use the "−−no−toc" option:

```
% htmldoc --no-toc ...
```

Disabling the Title Page

The title page is normally generated for all HTML, PostScript, and PDF output. To turn the title page off use the "−−no−title" option:

```
% htmldoc --no-title ...
```

Changing the Body (Background) Color

Use the "−−bodycolor" option to change the background color:

```
% htmldoc --bodycolor #RRGGBB ...
```

The color can be any primary color (black, red, green, yellow, blue, magenta, cyan, or white) or a specific red−green−blue value.

Changing the Body (Background) Image

Use the "−−bodyimage" option to change the background image:

```
% htmldoc --bodyimage filename ...
```

The image file can be any PNG, GIF, or JPEG image.

HTML−Specific Options

The following options apply to HTML output.

Changing the Navigation Bar Color

Use the "−−barcolor" option to match the navigation bar color to your logo image:

```
% htmldoc --logo logo.gif --barcolor #RRGGBB ...
```

The color can be any primary color (black, red, green, yellow, blue, magenta, cyan, or white) or a specific red−green−blue value.

PostScript−Specific Options

The following options apply to PostScript output.

Forcing Grayscale Output

To force all output to be in grayscale use the "−−gray" option:

```
% htmldoc --gray -f outfile.ps ...
```

This option is necessary for all B&W Level 1 PostScript printers.

Using JPEG Compression

To use JPEG compression for large images use the "−−jpeg" option:

```
% htmldoc --jpeg -f outfile.ps ...
```

The default JPEG quality is 90; to set a different quality use:

```
% htmldoc --jpeg=quality -f outfile.ps ...
```

where *quality* is the standard JPEG quality level from 1 to 100.

JPEG compression is not available on Level 1 PostScript printers.

Forcing Level 1 Output

To force Level 1 PostScript output use the "−t ps1" option:

```
% htmldoc -f outfile.ps -t ps1 ...
```

This option is necessary for all Level 1 PostScript printers.

Requesting Double−Sided Output

The "−−duplex" option specifies double−sided output:

```
% htmldoc --duplex -f outfile.ps ...
```

Note that this does not select duplexing on the printer but merely adjusts the formatting so that the left & right margins are swapped on the back side and chapters start on an odd−numbered page. You must still select duplexing in your printer driver or on the printer itself.

Setting the Page Size

The "−−size" option specifies the output page size:

```
% htmldoc --size letter ...
% htmldoc --size a4 ...
% htmldoc --size universal ...
% htmldoc --size WIDTHxHEIGHT ...
% htmldoc --size WIDTHxHEIGHTin ...
% htmldoc --size WIDTHxHEIGHTcm ...
% htmldoc --size WIDTHxHEIGHTmm ...
```

The "WIDTH" and "HEIGHT" arguments can be in points (no units specified), inches, centimeters, or millimeters. The default page size is Universal (8.27x11in or 210x279mm) which is the minimum of the US and European standard sizes (Letter and A4, respectively).

Note that this does not select a media size on the printer but merely adjusts the formatting so that the text and images appear within the given page area. You must still select the appropriate media size in your printer driver or on the printer itself.

Setting the Page Margins

The "−−left", "−−right", "−−top", and "−−bottom" options control the page margins of the output. The defaults are 1 inch (25mm) for the left and 0.5 inches (12mm) for the right, top, and bottom margins.

Setting the Default Font Typeface and Size

The default font size, spacing, and typefaces are controlled by the "−−fontsize", "−−fontspacing", "−−bodyfont", and "−−headingfont" options:

```
% htmldoc --fontsize 9.0 --fontspacing 2.0 ...
% htmldoc --bodyfont helvetica ...
% htmldoc --headingfont times ...
```

The typefaces for "−−bodyfont" and "−−headingfont" can be "courier", "times", or "helvetica".

Customizing the Page Headers and Footers

The "−−header" and "−−footer" options allow you to customize the headers and footers used for the document body. Each option requires a three character string that specifies the left, middle, and right fields:

Char	Description
.	A period indicates that the field should be blank.
t	A "t" indicates that the field should

	contain the document title.
h	An "h" indicates that the field should contain the current heading.
l	A lowercase L indicates that the field should contain the logo image.
1	The number 1 indicates that the field should contain the current page number in decimal format (1, 2, 3, ...)
i	A lowercase I indicates that the field should contain the current page number in lowercase roman numerals (i, ii, iii, ...)
I	An uppercase I indicates that the field should contain the current page number in uppercase roman numerals (I, II, III, ...)

The "−−tocheader" and "−−tocfooter" options control the header and footer on table−of−contents pages.

Setting the Header and Footer Font

The "−−headfootsize" and "−−headfootfont" options set the size and typeface of the font used for the page headers and footers:

```
% htmldoc --headfootsize 9.0 --headfootfont courier ...
```

PDF−Specific Options

The following options apply to PDF output.

Forcing Grayscale Output

To force all output to be in grayscale use the "−−gray" option:

```
% htmldoc --gray -f outfile.pdf ...
```

Using JPEG Compression

To use JPEG compression for large images use the "−−jpeg" option:

```
% htmldoc --jpeg -f outfile.pdf ...
```

The default JPEG quality is 90; to set a different quality use:

```
% htmldoc --jpeg=quality -f outfile.pdf ...
```

where *quality* is the standard JPEG quality level from 1 to 100.

Requesting Double−Sided Output

The "−−duplex" option specifies double−sided output:

```
% htmldoc --duplex -f outfile.pdf ...
```

Note that this does not select duplexing on the printer but merely adjusts the formatting so that the left & right margins are swapped on the back side and chapters start on an odd−numbered page. You must still select duplexing in your printer driver or on the printer itself.

Setting the Page Size

The "−−size" option specifies the output page size:

```
% htmldoc --size letter ...
% htmldoc --size a4 ...
% htmldoc --size universal ...
% htmldoc --size WIDTHxHEIGHT ...
% htmldoc --size WIDTHxHEIGHTin ...
% htmldoc --size WIDTHxHEIGHTcm ...
% htmldoc --size WIDTHxHEIGHTmm ...
```

The "WIDTH" and "HEIGHT" arguments can be in points (no units specified), inches, centimeters, or millimeters. The default page size is Universal (8.27x11in or 210x279mm) which is the minimum of the US and European standard sizes (Letter and A4, respectively).

Note that this does not select a media size on the printer but merely adjusts the formatting so that the text and images appear within the given page area. You must still select the appropriate media size in your printer driver or on the printer itself.

Setting the Page Margins

The "−−left", "−−right", "−−top", and "−−bottom" options control the page margins of the output. The defaults are 1 inch (25mm) for the left and 0.5 inches (12mm) for the right, top, and bottom margins.

Setting the Default Font Typeface and Size

The default font size, spacing, and typefaces are controlled by the "−−fontsize", "−−fontspacing", "−−bodyfont", and "−−headingfont" options:

```
% htmldoc --fontsize 9.0 --fontspacing 2.0 ...
% htmldoc --bodyfont helvetica ...
% htmldoc --headingfont times ...
```

The typefaces for "−−bodyfont" and "−−headingfont" can be "courier", "times", or "helvetica".

Customizing the Page Headers and Footers

The "−−header" and "−−footer" options allow you to customize the headers and footers used for the document body. Each option requires a three character string that specifies the left, middle, and right fields:

Char	Description
.	A period indicates that the field should be blank.
t	A "t" indicates that the field should contain the document title.
h	An "h" indicates that the field should contain the current heading.
l	A lowercase L indicates that the field should contain the logo image.
1	The number 1 indicates that the field should contain the current page number in decimal format (1, 2, 3, ...)
i	A lowercase I indicates that the field should contain the current page number in lowercase roman numerals (i, ii, iii, ...)
I	An uppercase I indicates that the field should contain the current page number in uppercase roman numerals (I, II, III, ...)

The "−−tocheader" and "−−tocfooter" options control the header and footer on table−of−contents pages.

Setting the Header and Footer Font

The "−−headfootsize" and "−−headfootfont" options set the size and typeface of the font used for the page headers and footers:

```
% htmldoc --headfootsize 9.0 --headfootfont courier ...
```

Disabling Document Compression

Normally each page in a PDF file is compressed using the Flate method (GZIP). Versions of Acrobat Reader prior to 3.0 do not understand Flate compression. To disable compression use the "−−no−compression" option.

A − Implementation Limits

HTMLDOC current has the following implementation limits:

- Generated PostScript/PDF pages: 5000
- Chapters: 100
- Headings: 10000
- Links: 20000
- Table columns: 20
- Table rows: 1000

This limits can be increased by changing the appropriate constant definition in the *config.h* header file.

B − GNU General Public License

GNU GENERAL PUBLIC LICENSE
Version 2, June 1991

GNU GENERAL PUBLIC LICENSE
TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program's source code as you receive it, in any

medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

> a. You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.
>
> b. You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.
>
> c. If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

> a. Accompany it with the complete corresponding machine−readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b. Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine−readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c. Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty−free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and "any later version", you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER

PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

END OF TERMS AND CONDITIONS