

Writing a big book—A first experience with \LaTeX

David Walden

This paper is in response to TUGboat editor Barbara Beeton's Editorial Wish List on the TUG web site for papers by non-experts for non-experts. I also thought my experience might be interesting because many other people have had to struggle with a book or journal publisher that has no interest in \LaTeX ; while my approach is quite ad hoc with many manual rather than computerized steps, such an approach can be (or at least seem to be) faster than seeking a fully computerized approach.

Section 1 describes why I decided to write a substantial book in \LaTeX , absent prior experience with \LaTeX , starting from a manuscript of a prior edition of the book prepared in Word, for a publisher with no knowledge in \LaTeX . My steps included converting the prior Word manuscript to \LaTeX (section 2), massively changing the manuscript in \LaTeX to produce a fully designed manuscript output in PDF format (section 3), and then converting the \LaTeX via HTML into Word files containing ASCII text for the publisher to input into QuarkXPress for production of the pages to be sent to the printer (section 4). I conclude with some reflections on my process (section 5).

1 Background

Several years ago, the publisher of a text book [1] I had co-authored asked for a second edition. The original book was 575 pages long. Because of advances in the field since the first edition's publication, the second edition would be significantly longer and quite different from the first edition. In effect, we would be writing a new 755 page book [2].¹

At about the same time the request came from my publisher, my frustration with composing large documents in Microsoft Word was reaching a peak. Because of new releases of Word and switches I made from using Word on early IBM PCs to using it on Macs to using it again on circa 2000 PCs running Microsoft Windows, I was having trouble accessing Word files I had created years before. I vowed never again to create a large document that had proprietary² formatting information as Word has; I wanted ASCII text with explicit markup commands,

¹ Throughout all of the steps described in what follows, my co-author Shoji Shiba worked collaboratively with me to create the content of the book; however, I did all the keyboarding and other computer-based work.

² Invisible and not able to be made visible, undocumented or unpublished, totally controlled by Microsoft.

so I could reprocess a document at any later time using a text editor and its macro capability or by writing a small text transformation program (for example, in Perl).

Also, I didn't know how to make Word do the following sorts of things (if it can do them at all):

- automatically adjust chapter, section, figure, etc., numbering as chapters, sections, figures, etc., are added, subtracted, and reordered;
- automatically order table of contents entries and insert page numbers;
- automatically adjust cross-references to chapters, sections, figures, etc., as necessary;
- support a file of include-file commands to permit working with various portions of the book at different times;
- allow global switching on and off of various manuscript components (such as marginal notes to myself about what still needed to be done);
- including EPS files of figures at specified points in the manuscript while avoiding awkward page breaks.

Therefore, I decided to learn and use \LaTeX which I had long known about but never used.³ My learning of \LaTeX was a bit hit and miss, as I started rewriting my book.⁴ Over time, I did the following:

- bought some books on \TeX and \LaTeX ;
- found www.tug.org and thus the Mi \TeX collection of software (see www.miktex.org);
- began watching the `comp.text.tex` newsgroup and thus stumbled across the series of lectures on \LaTeX and PDF by David Arnold of California's College of the Redwoods;⁵
- learned of WinEdt (www.winedt.com);
- configured WinEdt to run \LaTeX , run DVIPS, and Distill the result into PDF files.

2 Moving the first edition book into \LaTeX

I got the first edition Word files back from the publisher and repeatedly used WinEdt's Find and Replace commands to convert the Word file for each chapter into a \LaTeX file:

³ I had used programs such as RUNOFF (under the CTSS operating system), MRUNOFF (under TENEX), and TROFF-NROFF (under UNIX).

⁴ There were many more steps in learning \LaTeX and rewriting my book than I describe in this paper, including plenty of missteps. In this paper, I skip much confusing detail and simplify the ordering of steps a little to allow the reader to understand the main thread of my story. Any reader who wants more detail should ask me for it.

⁵ <http://online.redwoods.cc.ca.us/instruct/darnold/index.htm>.

- open and close smart double quotes from Word were changed to ‘ ‘ and ’ ’, and open and close smart single quotes were changed to ‘ and ’;
- em-dashes were changed to ---;
- hyphens between numbers were changed to --;
- footnotes (with the invisible, proprietary Word markup) were changed to use the `\footnote` command in L^AT_EX;
- the escape character `\` was inserted before `$`, `%`, etc.;
- ellipses (three periods in a row) were changed to the `\dots` command;
- chapter and section titles were enclosed in appropriate `\chapter` and `\section` commands;
- the command `\addcontentsline{toc}{chapter}{title}` was used to add table of contents lines without chapter numbers for the *title* of the five major parts of the book;
- italic text was enclosed in `\emph` commands, bold face text was enclosed in `\textbf` commands, various non-English letters were handled with a `\` and appropriate diacritical mark, etc.

The steps just listed took half an hour or so per file for each of the 20 or so files of the first edition book.

Next, I wrote some new L^AT_EX environments and commands, including the following:

- `blist`, `nlist`, and `dlist` environments for enumerated, numbered, and description lists with tighter inter-item spacing than the L^AT_EX default;⁶
- `\mnote{text}` command to put *text* to myself in the margin and `\CK` command to put a bold **CK** in the manuscript to highlight facts that needed checking;
- `\snfig`, `\snufig`, `\swsnfig`, etc., commands to include EPS figures in the text with the option of having them numbered and captioned, unnumbered and uncaptioned, sideways on a page, etc., and a similar set of commands to insert tables⁷ in the text. See Figure 1 for an example of the definition of one of these commands.⁸

⁶ After I finished the effort described here, I learned about the `memoir` class that includes support for tighter spacing of list items.

⁷ I used L^AT_EX’s `\tabular`, etc., features to format a few simple tables. Mostly, however, I already had tables or created tables as EPS files with Illustrator.

⁸ The macro for `\snfig` inserts the figure using the `graphics` package with parameter `dvips`, creates a label for it derived from the figure’s file name for cross-referencing from elsewhere in the book, and prints enough of the figure’s file

```

\newcommand{\snfig}[3]{ %scaled numbered figure
\begin{figure}[htbp]
\hfil\scalebox{#3}{%
\includegraphics{figures/fig#2.eps}%
}\hfil
\caption{\label{fig:#2}#1%
\texttt{\small[fig#2]}}%
}
\end{figure}
}

```

Figure 1: Definition of `\snfig`

Once these environments and commands were available,

- I converted the lists in the original manuscript to use the appropriate new environments by typing the necessary `\begin`, `\end`, and `\item` statements into each file of the manuscript.
- I converted absolute figure and table references from the first edition into references to labels, using `\ref` and `\pageref`.
- I converted absolute chapter and section references to `\ref` and `\pageref` references to appropriate labels I had inserted in the L^AT_EX versions of the chapter files.⁹

Finally, I compiled the L^AT_EX files and printed them out to compare them with the first edition book, looking for instances where I had missed something that needed to be converted to L^AT_EX.

3 Rewriting and updating the book

With the first edition fully converted to L^AT_EX as described in the previous section, it was time to move on to updating the book.

I drew additional text from other papers I and my co-author had written since the first edition of the book was published, doing the same brute force but fast enough conversions from Word to L^AT_EX. We wrote lots of new text that I typed into the L^AT_EX

name in square brackets on the page to which the figure floats so that I could find the figure’s file when I saw a figure needing a change.

The macro as shown here has extra new lines (so it fits within a single column) that were not in the macro as I actually used it.

⁹ I didn’t redefine the `\chapter` and `\section` commands to automatically produce appropriate labels because I didn’t want to take the time to learn how to do this and discover whether it was actually a good idea or not. Perhaps I should have. Generally I was focused on getting the new book written and sent to the publisher rather than on learning to be a L^AT_EX expert.

files. We created many new figures and tables that I input into Illustrator files.¹⁰

An important aspect of the book was a substantial number of real life case studies and examples. Thus, I decided to include a table of case studies (to go after the table of contents), listing the title of each case study and the page number on which it began. Rather than learn how to define a \LaTeX construct that worked like a table of contents such that a new entry was automatically added each time a new case study was added, I simply typed the case study name and a `\pageref` into a case-study-list file in the order of the case studies in the book. When I rearranged parts of chapters so the order of case studies changed, I had to do a parallel rearrangement of the lines of the case-study-list file. It was easy to see when rearrangement of the case-study-list file was needed by the out-of-order page numbers.

Eventually, I was ready to show a first draft of the new book to the development editor from the publisher's staff. I used the command

```
\setlength{\parskip}{.24in}
```

to (sort of) produce a double-spaced copy of the manuscript for the development editor to review.¹¹

Based on the development editor's comments, I further revised and reordered the manuscript. In particular, the editor asked that a single bibliography be created as part of the backmatter of the book to which cross-references could be made from the main text, a significant change from the first edition which included citations in the end notes for each chapter. Thus, I began to use \BibTeX , creating a `.bib` file of all bibliographical items and replacing citations in the text with appropriate uses of the `\cite` command. If there is one thing I wish I'd done differently in my work on the new book, it is to use \BibTeX and a `.bib` file from the start. Note only did it eliminate lots of chapter end notes; it also provided a single extensive bibliography, arranged alphabetically by author, which has been a valuable addition to the book.

In time,¹² a final manuscript was available.

¹⁰ The publisher did the first edition figure finish work in Illustrator from MacDraw Pro figures I submitted with the Word manuscript. Therefore, I used Illustrator to draw new figures and tables and to modify old figures and tables. Altogether, the finished book had hundreds of figures and tables in EPS format.

¹¹ I always did my own proofreading and reviews of what I was writing using a single-spaced copy of the manuscript because it helped me better picture the final book I was aiming toward.

¹² The total elapsed time to rewrite the book was about 18 months, most of which was involved with content rather than formatting.

Again, I provided double-spaced pages for copy editing. However, by this time, the publication deadline was looming, so as I finished the final draft of each chapter, I sent it as an email attachment to the copy editor, who did his work and then Fedexed red-marked pages back to me from which I made appropriate changes in the \LaTeX files.

4 Producing the final book

With a finished manuscript in hand, it was time to refine the book design so the publisher could see precisely what I wanted. This required using the `fancyheadings` package to set up page headers and footers, using the `titlesec` package and then making a tiny change to a private copy of the `titlesec` style to adjust chapter headings, and slightly redefining (in the \LaTeX preamble) \LaTeX 's book class (`\@makecaption` and `\thechapter` as used in table and figure captions) to get captions to look as I wanted. I learned how to do these redefinitions by sending a query to `comp.text.tex`.

It was also time to discover the format the publisher wanted the final manuscript in — a topic I had been avoiding because I didn't want to hear any back pressure against my use of \LaTeX for the book rewrite. I offered some choices to the publisher's production manager for the book:

1. I could finish the job in \LaTeX (getting rid of strange page breaks, etc.) and deliver a ready-to-print PostScript file for the book.
2. I could convert the manuscript back to Word, fully formatted (I didn't know how I would accomplish this with other than *lots* of manual work).
3. I could do something in between 1 and 2.

The publisher's layout person and I communicated. He stated that he would be using QuarkXPress to lay out the book. Once he understood what I had in \LaTeX , he asked me to provide him with ASCII text in Word files with no formatting: he wanted Word files because in his experience special characters such as non-English vowels including diacritical marks translate best into XPress from Word; he wanted no formatting so he wouldn't have to remove anything unnecessary before adding formatting in XPress.

My problem, therefore, was how to convert the *output* of \LaTeX (with all the benefits of \LaTeX 's chapter, section, figure, citation, etc., numbering and cross-referencing) into Word files (without any \LaTeX markup, \LaTeX -produced hyphenation, etc.). I decided that moving from files with \LaTeX markup through HTML produced by \LaTeX to ASCII in Word

files would result in what the publisher's layout person needed.

Therefore, I spent a few days (*just a few days*) doing the following:

1. Sent the layout person a single-spaced copy of the fully designed and fully formatted book out of L^AT_EX so he had that to guide him as he did his work in XPress.
2. Replaced the `\footnote` commands in all 30 or so L^AT_EX files with instances of a newly defined `\enote` command that produced end notes following each chapter, using the `endnotes` package.¹³
3. Modified my figure- and table-producing commands (like `\snfig` shown in Figure 1) so they dropped the marginal notes indicating the figure file names, did not actually insert the EPS file, and instead just included the figure number and caption on its own line at an appropriate place in the manuscript along with the file name of the figure or table. The locations of the modified figure- and table-producing commands were always immediately after the the paragraph of first reference to the figure or table. Thus, the person doing the layout had the information necessary to place each EPS figure or table with its number and caption at an appropriate place on an appropriate page.
4. Added the command `\setlength{\defaulthyphenchar}{-1}` in the preamble to turn off hyphenation to avoid artificial within-word breaks within paragraphs.
5. Bought a copy of V_TE_X, the advertising for which claimed it to be the best program for generating HTML from L^AT_EX because it generated HTML from the L^AT_EX itself and not from DVI output of L^AT_EX.¹⁴
6. Changed a few small items to conform with V_TE_X, e.g., from using the `graphics` package I had been using with M_IK_TE_X to using the `graphicx` package.
7. Got a small modification to V_TE_X from MicroPress Inc. so that V_TE_X generated no extra

¹³ While I was writing the new book I used bottom-of-the-page footnotes because of their proximity to the text they augmented. However, my publisher preferred all such notes to be at the end of the book as a collection of end notes, to avoid having many footnotes on text pages that might make the book look quite technical and, thus, less popular. We compromised on having end-of-chapter end notes.

I made a couple of tiny modifications to a private copy of the `endnotes.sty` file in order to format the chapter end notes as I wanted them to be.

¹⁴ I have heard that other software packages, such as `latex2html`, also generate HTML directly from L^AT_EX.

characters in its HTML output (e.g., no gratuitous spaces surrounding bibliographic reference numbers).

8. Processed each L^AT_EX chapter individually (using `\includeonly`) through V_TE_X to produce an HTML file for the chapter.
9. For each chapter, selected and copied (using control-C under Windows) the entire text of the HTML output file and pasted it (using control-V) into a new WinEdt file.
10. Configured WinEdt to have essentially infinitely long lines of text and touched each paragraph with the mouse and keyboard so that each paragraph became one line. In other words, I removed intra-paragraph new lines that would interrupt the flow of text on to pages in XPress.
11. Selected and copied (again using Windows' control-C command) all the text for the chapter in WinEdt and pasted it into an empty file in Word (using control-V). Somehow moving all the text from the HTML file to WinEdt and then to Word discarded all of the HTML markup commands and left only ASCII text of the manuscript in the Word file, which wouldn't have happened with a direct copy from the HTML file to Word.
12. Reviewed the Word file for anything that might confuse the layout person and clarified it.¹⁵
13. Sent each Word file to the layout person to cut-and-paste into XPress and there to format it for printing.

The layout person did the layout of the whole book and then sent it to a proofreader. The proofreader reviewed the entire set of page proofs, marked them where corrections were apparently needed, and sent them to me. I reviewed the proofreader's notes, changed my L^AT_EX files appropriately (to keep them in sync with the book as published), and sent the page proofs back to the layout person with instructions to ignore the few changes from the proofreader I didn't want made. Somewhere around this same time, the page proofs went to the indexer who created the index and sent it to the layout person to include at the end of the book.¹⁶

¹⁵ Somewhere along the line, about this time I formatted the output of `\tabular` commands so the layout person would recognize it and format the content with XPress.

¹⁶ I may misremember some of the ordering of the passing of nearly final manuscript from person to person. In any case, what I have described in this paragraph is what, in effect, happened (including no use of L^AT_EX's indexing capabilities).

5 Reflections

5.1 On using L^AT_EX

I surely could have been smarter about the way I used L^AT_EX. I feel like I could have made more use of redefining environments and commands. I could have done a better job of making it possible to change one environment or command definition to change the behavior (e.g., switching it off or on) of all instances of use of the environment or command. I learned to do almost nothing with fonts other than use the built-in ones and switch among their standard modes (and, since my book had no math in it, I learned nothing about L^AT_EX's math layout capabilities). I am sure there were "right ways" to do things that I did with the first string of commands or characters that I stumbled across that achieved the effect I wanted.

I will welcome suggestions from readers about smarter use of L^AT_EX.

5.2 On writing a book in L^AT_EX

The choice to rewrite the new edition of the book in L^AT_EX is one I would make again. Conceivably it would have taken less overall effort to do it all in Word.¹⁷ However, I would not have been as happy as I was seeing and flexibly reorganizing and rewriting the book-like manuscript with embedded figures I had with L^AT_EX. (Also, I benefited by learning a lot about L^AT_EX.) The final conversion from ASCII text in L^AT_EX files to ASCII text in Word files for input to the layout person and XPress was a simple and quick enough step, given the V_TE_X capability to generate HTML without extra HTML formatting.

I will welcome critiques and suggestions for a better overall strategy for creating the design and manuscript of a big book in a way that provides me as author with lots of flexibility and a thorough image of the final product while at the same time working with a publisher apparently unacquainted with the likes of L^AT_EX and not used to authors so willing to get deeply into the computer aspects of creating the files to be sent to the printer.

¹⁷ It might also have been more efficient for me to do the entire rewrite of the manuscript in XPress for continuity with the final layout work by the publisher. However, at the beginning of the project the publisher's art department was adamant that an author could not be allowed near XPress—that was a domain reserved for the publisher's layout people.

Acknowledgments

The publisher's production editor, Michael Ryder, and composition person, William Brunson, cooperated completely in my effort to convert the L^AT_EX files into files the publisher needed to have the book printed.

Michael Vulis of MicroPress Inc. quickly provided a needed small change to V_TE_X. Peter Flynn answered my query to the `comp.text.tex` newsgroup regarding making captions look as I wanted them to. I also used many other resources available to the T_EX community, for example, at CTAN and on the T_EX Live CDs that come with membership in the T_EX Users Group (along with a subscription to *TUGboat*).

Of course, many other people helped with the non-L^AT_EX-related aspects of writing the book. They are acknowledged in the book itself.

I appreciate the guidance for preparing this paper for publication that I received from Barbara Beeton, Karl Berry, and from the anonymous reviewer.

References

- [1] Shoji Shiba, Alan Graham, and David Walden. *A New American TQM*. Productivity Press, Portland, OR, 1993.
- [2] Shoji Shiba and David Walden. *Four Practical Revolutions in Management*. Productivity Press, Portland, OR, 2001.

Author's biography

Over a thirty-plus year career in high tech computer R&D, David Walden worked successively as a technical contributor, a technical manager, and a general manager. Most of this time, he was with Bolt Beranek and Newman Inc. of Cambridge, MA, where he was part of the original ARPANET development team and involved in many other early Internet activities. He has done much writing on technical, management, and other topics.

◇ David Walden
 12 Linden Road
 E. Sandwich, MA 02537
dave@walden-family.com
<http://www.walden-family.com>