The TeX tuneup of 2014

Donald Knuth

If you ask the Wayback Machine to take you back to the home page

```
http://www-cs-faculty.
        stanford.edu/~knuth/abcde.html
```

of *The TeXbook* and my other books on *Computers & Typesetting*, as that page existed on 16 January 1999, you'll find the following remarks:

> I still take full responsibility for the master sources of TeX, METAFONT, and Computer Modern. Therefore I periodically take a few days off from my current projects and look at all of the accumulated bug reports. This happened most recently in 1992, 1993, 1995, and 1998; following this pattern, I intend to check on purported bugs again in the years 2002, 2007, 2013, 2020, etc. The intervals between such maintenance periods are increasing, because the systems have been converging to an error-free state.

And if you fast-forward nine more years, you can find a *TUGboat* article called "The TeX tuneup of 2008" [4], which describes the changes that were made to TeX and its companion systems based on the comments from users that were received during the years 2003, 2004, 2005, 2006, and 2007. That article ended as follows:

> So now I send best wishes to the whole TeX community, as I leave for vacation to the land of *TAOCP* — until 31 December 2013. Au revoir!

Hello again, dear friends, *allô*! Here is the sequel.

On 31 December 2013, Barbara Beeton duly forwarded to me a well-organized collection of materials covering more than two dozen potentially troublesome topics that had been submitted for consideration during the years 2008, 2009, 2010, 2011, 2012, and 2013. This was the residue of hundreds of items that had been carefully filtered by a team of expert volunteers, who had worked hard to minimize the effort that I would need to devote to this project. (I can't possibly thank all the volunteers individually; but Donald Arseneau, Karl Berry, Peter Breitenlohner, and Bogusław Jackowski deserve particular commendation.)

As in 2008, both TeX and METAFONT have changed slightly and gained new digits in their version numbers. But again, the changes are essentially invisible. I can't resist quoting another paragraph from [4], because it reflects my unwavering philosophy (see [3]):

> The index to *Digital Typography* lists eleven pages where the importance of stability is stressed, and I urge all maintainers of TeX and METAFONT to read them again every few years. Any object of nontrivial complexity is non-optimum, in the sense that it can be improved in some way (while still remaining non-optimum); therefore there's always a reason to change anything that isn't trivial. But one of TeX's principal advantages is the fact that it does not change — except for serious flaws whose correction is unlikely to affect more than a very tiny number of archival documents.

Users can rest assured that I haven't "broken" anything in this round of improvements. Everyone can upgrade at their convenience.

TeX Version 3.14159265

Let's get down to specifics. The new version of TeX differs from the old only with respect to the "null control sequence" `\csname\endcsname`, which has been a legal construct since version 0.8 (November 1982) although almost nobody uses it. Oleg Bulatov noticed in September 2008 that TeX's `\message` operation has curiously inconsistent behavior: Suppose you say

```
\def\\#1{\message{#1bar}}
\def\surprise{wunder}
\let\foo=!
```

(for example). Then

```
\\\surprise                  gives wunderbar
\\\over                      gives \over bar
\\\foo                       gives \foo bar
\\{\csname 6\endcsname}      gives \6bar
\\{\csname fu\endcsname}     gives \fu bar
```

as messages on your terminal and in your log file. But '`\\{\csname\endcsname}`' unfortunately gives

```
\csname\endcsnamebar
```

because I forgot to insert a space when I coded this part of the *print_cs* routine (see [B], §262). So Oleg has won a check for $327.68 [1]. Of course I hope that this turns out to be the "historic" final bug in TeX. (It's the 947th; see [3], page 662.)

Henceforth '`\\{\csname\endcsname}`' will give

```
\csname\endcsname bar
```

and everybody will be happy. This corrected behavior does not simply affect TeX's messages; the name of a control sequence can also get into documents, for example via `\write` or `\meaning`. But the change surely won't ruin your archived works.

METAFONT Version 2.7182818

The historic final (I hope) bug in METAFONT was discovered during June 2008 by the longstanding TeX contributor Eberhard Mattes. The error that he brought to light is easier to describe than the TeX error discussed above, but it was much more subtle to detect: Whenever previous versions of METAFONT have transformed **pencircle** into an axially symmetric pen whose polygon has no point on the x-axis, the algorithm in §536 of [D] has "leaked memory," by forgetting to reclaim seven words that had been allocated for the omitted point. This happened, for instance, with one of the pens in exercise 16.2 of [C], and in my original `TRAP` test [2] for METAFONT; so I should have discovered the problem long ago. Eberhard noticed that the METAFONT program

```
pen p;
forever:  showstats;
  p := pencircle scaled 1.4; endfor
```

would abort with METAFONT's capacity exceeded—although it did take quite awhile to overflow 3 million words of memory on my current home system—and he also figured out how to cure the problem. For this he amply deserves his new reward in [1].

Computer Modern

No changes have been made to the Computer Modern fonts of 2008, although I did delete a few bytes of redundant source code and alter two names.

John Bowman noticed a tiny bump that appears near the top right serif when an italic 'K' is greatly magnified, and Jacko discovered the underlying reason: Part of the stroke of this slanted letter is drawn with a circular pen, but it joins up with outlines that are slanted (hence not true circles). The same tiny bumps can therefore by observed also in various other italic and slanted letters, such as A, V, W, X, Y, when enlarged.

But those bumps are even less visible than the mispositioned bulbs that I discussed in [4]. And in fact I've even become somewhat fond of such little glitches, now that I've been learning to appreciate the Japanese concept of *wabi-sabi*.

Thus I've decided that the Computer Modern fonts are to be forever frozen in their present form, especially now that the definitive description in the latest printing of [E] has become available.

TeXware and METAFONTware

I made minor updates to the master `web` files for five other programs, namely `gftopk`, `pltotf`, `tftopl`, `vftovp`, and `vptovf`, in order to make them more robust in the presence of weird input files. (These changes had in fact already been made in recent editions of TeX Live; now they are in some sense "official.") Here is a current list of all the `web` files for which I have traditionally been responsible:

name	current version	date
`dvitype.web`	3.6	December 1995
`gftodvi.web`	3.0	October 1989
`gftopk.web`	2.4	January 2014
`gftype.web`	3.1	March 1991
`mf.web`	2.7182818	January 2014
`mft.web`	2.0	October 1989
`pltotf.web`	3.6	January 2014
`pooltype.web`	3.0	September 1989
`tangle.web`	4.5	December 2002
`tex.web`	3.14159265	January 2014
`tftopl.web`	3.3	January 2014
`vftovp.web`	1.4	January 2014
`vptovf.web`	1.6	January 2014
`weave.web`	4.4	January 1992

Typographic errors and other blunders

So far I've only been discussing potential anomalies in the software. But of course people have also reported problematic aspects of the documentation—which may actually be the hardest thing to get right. Even *The TeXbook* [A], which has been under intense scrutiny for more than thirty years, was not free of hitherto-unperceived defects.

Altogether I made corrections to each of [A], [B], [C], [D], and [E], enough to represent $23.68 in eleven new reward checks. The most significant of these changes can be seen from the home page cited above, if you click to get the PDF errata file and scan for corrections dated in 2014.

The master sources

The backbone of the TeX system, for the past 25 years or so, has been a collection of 178 files, mostly with names of the forms `*.web`, `*.tex`, and `*.mf`. These files contain almost exactly 7 megabytes altogether; and the new changes have altered about 3500 of those bytes. Thus it appears that the TeX system was 99.95% correct in 2008, if it is 100% correct today.

The master files, together with a bunch of errata files that document past history, can be downloaded from the ftp server `cs.stanford.edu`, which accepts '`anonymous`' as a login name. They're collected together in a single compressed file

`pub/tex/tex14.tar.gz`,

which you can compare if you like to the older files `pub/tex/tex08.tar.gz`, `pub/tex/tex03.tar.gz`. The latest versions of individual files can of course also be found in the CTAN archive.

As I did in [4], I'll mention here the names of all files that have changed in some way during the latest go-round:

`tex/texbook.tex` % source file for [A]

`tex/tex.web` % master file for TeX in Pascal

`tex/trip.fot` % torture test terminal output

`tex/tripin.log` % torture test first log file

`tex/trip.log` % torture test second log file

`tex/trip.typ` % torture test output of DVItype

`texware/pltotf.web` % master file for PLTOTF

`texware/tftopl.web` % master file for TFTOPL

`mf/mfbook.tex` % source file for [C]

`mf/mf.web` % master file for METAFONT in Pascal

`mf/trap.fot` % torture test terminal output

`mf/trapin.log` % torture test first log file

`mf/trap.log` % torture test second log file

`mf/trap.typ` % torture test output of DVItype

`mfware/gftopk.web` % master file for GFTOPK

`cm/romanu.mf` % master file for Computer Modern Roman uppercase

`cm/symbol.mf` % master file for Computer Modern Roman symbols

`etc/vftovp.web` % master file for VFTOVP

`etc/vptovf.web` % master file for VPTOVF

`lib/manmac.tex` % macros for [A] and [C]

`errata/errata.nine` % changes to [A] between 1992 and 1996

`errata/errata.tex` % changes to [A]–[E] since 2001

`errata/tex82.bug` % changes to `tex.web`

`errata/errorlog.tex` % one-per-line annotated summaries of those changes

`errata/mf84.bug` % changes to `mf.web`

(Notice that the basic macro files for plain vanilla TeX and plain vanilla METAFONT, `lib/plain.tex` and `lib/plain.mf`, remain unchanged.)

Questions and answers

Barbara also asked me to answer three questions, which she said "keep coming up in various forums," so that she could point people to the answers if those questions come up again.

(1) How long did it take to typeset *The TeXbook* in the 80s, and how long does it take today?

This question is a bit strange, because anybody who tries to apply TeX to the file `texbook.tex` immediately gets the message '`~\.{This manual is copyrighted and should not be TeXed`', repeated endlessly. Therefore the running time to typeset *The TeXbook* has always been infinite.

On the other hand, I myself have to generate new printings every now and then; and I have a favorite way to get around the booby trap by first typing '`19`' and then typing some other special codes. (I also realize that unscrupulous people might even try to change `texbook.tex`, although that is strictly forbidden. The source code is intended to be *examined*, if desired, but not *executed* or modified except by its author.)

Unfortunately I don't think I ever noted down the running time in the 80s, so I can't give a definitive answer to the question. My recollection is that the entire book took maybe 20 minutes on Stanford's PDP10 mainframe (shared with other users). There was a noticeable slowdown on certain pages — such as page 218, when prime numbers are computed the hard way.

My colleague David Fuchs used *The TeXbook* as a benchmark in 1986, when he was developing MicroTeX (the first version of TeX to run on an IBM PC). A few days ago I asked him if he could remember its speed. He replied that, like me, he had no firm memory of those days, except that MicroTeX could do several pages per minute; and he guessed that it had taken roughly an hour to complete the whole *TeXbook*. His estimate seems right, because *The TeXbook* has nearly 500 pages.

Today, on my home computer (a 3.6 GHz Xeon with 10 MB cache), TeX transforms `texbook.tex` to `texbook.dvi` in 0.3 seconds.

(2) If you were designing TeX today, would you still use `\over` and friends, rather than something like `\frac{...}{...}`, when the latter would avoid the necessity of `\mathchoice` and `\mathpalette`?

This question, from `tex.stackexchange.com`, also quoted from page 151 of [A]:

> `\mathchoice` is somewhat expensive in terms of time and space, and you should use it only when you're willing to pay the price.

And well, I guess that quote implies my answer. For I was clearly willing to pay the price in 1982, so I'm certainly willing to pay zero today!

I suppose there are some people in the world who prefer expressions like '$\mathrm{sum}(2,3)$' to '$2+3$'; but I'm certainly not among them. Ever since TeX was born, I've been enormously pleased by the ability to write '`2\over3`' or '`n\choose k`' or '`p\atop q`' or $\cdots$, instead of being forced to write something like '`frac{2}{3}`' that would have distracted my attention from the task at hand.

The questioner seems to want to place burdens on all users, rather than on the backs of a few macro-developers.

(3) Why is the default rule thickness 0.4 points?

One of the very first things I did when designing TeX was to choose several publications that represented the highest standards of excellence in mathematical typesetting, and to "reverse engineer" them by making careful measurements of those fine works. (See [3], page 620.) The thickness of rules in *The Art of Computer Programming* was definitive for me. I also knew that Belfast Universities Press was using that value in its typesetting of mathematical journals in 1977.

This question, however, is related to the one sore point with respect to which I wish that I could turn back the clock and redesign TeX from scratch: The actual default rule thickness in TeX is not *exactly* 0.4 printer's points; it is exactly 26214 *scaled* points, where there are 65536 scaled points to every printer's point. Thus the default rule thickness is actually 0.399993896484375 points.

I made the foolish mistake of using binary fractions internally, while providing approximate decimal equivalents in the user interface. I should have defined a scaled point to be 1/100000 of a printer's point, thereby making internal and external representations coincide. This anomaly, which is discussed further in [5], is the only real regret that I have today about TeX's original design.

Conclusion

The TeX family of programs seems to be healthy as it continues to approach perfection. Volunteers have been stalwart contributors to this success in optimum ways. Stay tuned for The TeX Tuneup of 2021!

References

[1] The Bank of San Serriffe, account balances. See `http://www-cs-faculty.stanford.edu/~knuth/boss.html` (accessed January 2014).

[2] Donald E. Knuth, *A torture test for METAFONT*. Stanford Computer Science Report 1095 (Stanford, California: Stanford University Computer Science Department, January 1986), 78 pages.

[3] Donald E. Knuth, *Digital Typography* (Stanford, California: Center for the Study of Language and Information, 1999), xvi + 685 pages. CSLI Lecture Notes, no. 78. The second printing (2012) contains numerous corrections.

[4] Donald Knuth, "The TeX tuneup of 2008," *TUGboat* **29** (2008), 233–238. `http://tug.org/TUGboat/tb29-2/tb92knut.pdf`.

[5] Donald E. Knuth, "An earthshaking announcement." *TUGboat* **31** (2010), 121–124. `http://tug.org/TUGboat/tb33-3/tb105knut.pdf`.

[A] Donald E. Knuth, *The TeXbook* (Reading, Mass.: Addison–Wesley, 1984), x + 483 pages. Also published as *Computers & Typesetting*, Volume A. Currently in its 34th printing (paperback) and 19th printing (hardcover).

[B] Donald E. Knuth, *Computers & Typesetting*, Volume B, *TeX: The Program* (Reading, Mass.: Addison–Wesley, 1986), xvi + 594 pages. Currently in its 9th printing (hardcover).

[C] Donald E. Knuth, *The METAFONTbook* (Reading, Mass.: Addison–Wesley, 1986), xii + 361 pages. Also published as *Computers & Typesetting*, Volume C. Currently in its 12th printing (paperback) and 8th printing (hardcover).

[D] Donald E. Knuth, *Computers & Typesetting*, Volume D, *METAFONT: The Program* (Reading, Mass.: Addison–Wesley, 1986), xvi + 560 pages. Currently in its 6th printing (hardcover).

[E] Donald E. Knuth, *Computers & Typesetting*, Volume E, *Computer Modern Typefaces* (Reading, Mass.: Addison–Wesley, 1986), xvi + 588 pages. Currently in its 7th printing (hardcover).

◇ Donald Knuth
`http://www-cs-faculty.stanford.edu/~knuth`