

Treasurer's Report

(Continued from preceding page)

Notes:

- The 1983 budget column is identical to that published in TUGboat Vol. 4, #1. All expense figures include an AMS overhead charge of 18%.
1. There are 738 memberships/subscriptions: 188 foreign, including Canada and Mexico; 558, U. S. Beginning in 1984, foreign air mail postage is included in membership/subscription fee.
 2. 55 copies of reprints of Max Diaz's "Fácil T_EX" have been sold.
 3. TUG has 44 institutional members, listed on the inside cover of this issue. 20 - educational; 24 - non-educational.
 4. 84 individuals attended Michael Spivak's "Introductory AMS-T_EX82 Users Course" and 135 members participated in the summer meeting conducted at Stanford University, July 11-15, 1983. Representatives from Autologic, Hewlett-Packard, Imagen and Quality Micro Systems gave presentations. A second meeting on the East Coast is planned for 1984.
 5. Support is budgeted for attendance at one meeting of ANSI X3J6.
 6. Advertising of TUG and the TUG Meeting/Course was accomplished through a news release to 19 trade publications, several of which are known to have published the notice, in addition to direct mailings to members and former members.
 7. While TUG was becoming established during 1981 and 1982, the American Mathematical Society made available the services of Ray Goucher at no charge. He manages all the administrative details associated with TUG, to include daily income/expense accounting, budgeting/treasurer's reports, coordination of all aspects of meeting preparations/accounting, publicity, advertising, in addition to numerous other details. He was appointed TUG Business Manager at the Stanford Meeting in July.
 8. Money available to the Finance Committee to subsidize travel and membership fees for individuals when appropriate.
 9. Postage/express charges, telephone tolls and supplies, plus programmer and clerical services not associated with production of TUGboat.

Respectfully submitted,
Samuel B. Whidden, Treasurer

* * * * *

REPORT ON ANSI X3J6

Lynne Price

The American National Standards Institute Technical Committee ANSI X3J6 on Computer Language for the Processing of Text meets concurrently with the International Standardization Organization's Expert Group (ISO/TC 97/SC 5/EG CLPT) on the same topic. As described in TUGboat Vol. 3, No. 1, the committee's charter is to define a standard language for tasks such as text editing, text formatting, and generalized markup. Availability of a standard will promote the ease with which document source files can be moved from site to site and will reduce retraining of individual users who transfer (even temporarily) to a new system. The sixth working draft of the language specification was submitted for comment to ISO/TC 97/SC 5 (SC 5 is the subcommittee on Programming Languages of TC 97, the technical committee on Computers and Information Processing) on June 15. The comment period extends through September 15 and the following draft is expected early next year.

The standard is divided into a series of parts. The parts are distributed as separate documents so that individuals who wish to use only part of the material may conveniently do so. The parts are enumerated in the following table, where an asterisk marks parts that were included in the June version:

- * 1 General
- * 2 Vocabulary
- * 3 Programming Language
- * 4 Entry and Editing Functions
- * 5 Formatting and Composition Functions
- * 6 Document Markup Metalanguage
- 7 Markup Support Functions
- 8 Binding to the Graphical Kernel System
- 9 Application to "What You See is What You Get" (WYSIWYG) Processing
- 10 Registration Procedures

Of the six parts so far submitted to SC 5, Parts One, Two, Five and Six are the most polished. Part Three, which is not yet completed, describes an interpretive language geared toward text processing. While the editing, formatting, and markup functions described in other parts can be implemented in the language described therein, the implementation language is not dictated by the standard. The language of Part Three does provide the end user with the ability to build macros of editing and formatting commands. The current version of Part Four (Entry and Editing Functions) is an expression of a possible philosophy. It notes that while opinions on text editor commands and syntax are highly individual, most sophisticated features are

built from a very small number of editing primitives. It suggests therefore that the standard define any text editor implemented in the language of Part Three to be compliant and that an Appendix to the document contain examples of such definitions for typical line, character, and screen editors. These examples would suggest to users of related editors how the listed code could be modified to incorporate each user's preferred enhancements. Part Ten will describe procedures for registering items such as document styles, markup conventions, and formatting macros. Although registered items will not be standardized, they will be available to all users who wish to access them. Users at one site can thereby take advantage of work done by individuals at another location.

Copies of the current working draft can be obtained from

Charles D. Card
Sperry Corporation
M.S. C1-NE10
Blue Bell, PA 19424

In addition, visitors are welcome to observe and participate in the meetings. The next sessions will take place October 24–28 in Detroit, January 23–27 in Anaheim, and May 14–18 in Phoenix. TUG's X3J6 liaison is Lynne Price, who will happily transmit feedback from TUG members to the committee. TUG members are of course encouraged to contact her for more information.

* * * * *

Software

* * * * *

A NOTE ON HYPHENATION

Donald Knuth

Some people occasionally write to me about hyphenations that T_EX finds, because T_EX doesn't always match the way their own dictionaries do it. In almost all cases, such discrepancies prove to be unavoidable, because different dictionaries don't agree with each other.

Consider, for example, the word "process." T_EX hyphenates 'pro-cess', in accordance with *Webster's Third*, while many dictionaries say 'process'. I don't believe T_EX does anything wrong here; indeed, I would never like to see 'proc-' at the end of one line and 'ess' at the beginning of the next, since I would probably have already pronounced the word wrong in my mind before my eyes reached the second line.

Another interesting case is "performance." Here *Webster's Third* and *American Heritage*, etc., say 'per-form-ance', but T_EX says 'per-for-mance'. This case is interesting because it turns out that *Webster's New Collegiate*—published many years after the infamous *Third*—also says 'per-for-mance'; so does *Random House Unabridged*. The latter hyphenation is evidently more consistent with other words of English, since T_EX's patterns are based on a large mass of data, so here we see a trend in dictionaries to be more uniform.

So far I have run across only one improperly hyphenated word, in thousands of test pages: 'exam-smanship'. But I wasn't too upset, because I deserved such a fate after making up that word.

Bob Filman has also shown me the very unfortunate 'Di-jkstra'; there's a case where many T_EX users will want an entry in their exception dictionaries.

I think it would be useful to have a catalog of desirable hyphenation exceptions maintained somehow in TUGboat; let me begin this with its first entry, 'Dijk-stra'. Let me also beg readers not to contribute further entries unless they are sure that all of the standard authorities disagree with T_EX's hyphens. (Sometimes we have found that *Webster's* is not as good as others, but we usually have followed it.)

And one more point: If any computer center decides to preload different exceptions from those in plain T_EX (i.e., in the file HYPHEN.TEX), the changed exceptions should not under any circumstances be put into HYPHEN.TEX or PLAIN.TEX. All local changes should go into a separate file, so that T_EX will still produce identical results on all machines. You can run your program elsewhere by simply sending the file of local changes. In fact, I recommend not preloading those changes, but rather assuming that individual users will have their own favorite collection of updates to the standard format files.

* * * * *

Editor's note: At the TUG meeting, Don gave some interesting statistics on the performance of the hyphenation algorithm in T_EX82: For the 676 most common English words, hyphenation is 100% correct. And 89.7% of all English words are hyphenated correctly. So among the remaining 10% there must be a few words that might show up in a T_EX file. The editor of TUGboat will be happy to keep a list.