

TEX for 30,000

JAMES HASKELL, WALLY DESCHENE, AND ALAN STOLLEIS

James Haskell
Director of Information Systems
USDA—Forest Service
Intermountain Region
324 25th Street
Ogden, Utah 84405

Wally Deschene
Computer Specialist
USDA—Forest Service
Intermountain Fire Sciences laboratory
P.O. Box 8089
Missoula, MT 59807

Alan Stolleis
Research Assistant
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
alan@cssun.tamu.edu

ABSTRACT

Quality documents are an integral part of the Forest Service mission. These documents include reports to the public, the Congress and other oversight agencies. The integration of TEX into our existing word processing system (Data General's Comprehensive Electronic Office — CEO) meets our goal of providing a consistent user interface and does not require all employees to learn a new system.

1. Forest Service Organization and Mission

The Forest Service, which is the largest agency in the U.S. Department of Agriculture, is charged with the responsibility of managing 191 million acres of National Forest System lands. Included on these lands are more than 128,000 miles of streams and two million acres of lakes and reservoirs. Each year, National Forests host more than 40 percent of all outdoor recreation in the country. Forest Service programs include areas such as fire prevention, fire suppression, range, water, air, timber, mineral leasing, oil and gas leasing, wildlife, recreation, engineering, forest pest management, and an extensive research program.

To carry out these programs, the Forest Service employs nearly 30,000 people in 9 Regional Offices, 123 National Forest Supervisor Offices, 653 Ranger Districts, approximately 50 research locations along with the National Office in Washington, D.C. The entire workforce is distributed over 45 states.

2. Forest Service Technology Environment

Beginning in 1983, the Forest Service embarked upon the installation of a comprehensive nation-wide distributed processing system. The objectives of this system were to meet increasing demands placed on the agency for timely and accurate information, and to provide a consistent office automation environment throughout the Forest Service.

The computer hardware acquired includes more than 900 Data General MV/family super minicomputers. These systems communicate with each other and with mainframe systems at USDA National

Computer Centers at New Orleans and Fort Collins over an X.25 public packet switched network. Included in the network are over 200 remote satellite locations, data channels on Forest Service-owned microwave channels and other local area networks.

The office automation software being used is Data General's Comprehensive Electronic Office (CEO). The systems are used for electronic mail, document preparation, electronic filing, spreadsheets, database creation and maintenance, graphics, most business transactions, and a variety of other uses that even include business management activities at forest fire camps. The CEO word processing system is the backbone and is a typical word processing system, including features such as format rulers, centering, bold, underline, tabs, indents, footnotes, super/subscripts, and more.

Today, nearly every one of the 30,000 employees has access to a terminal somewhere in the system to conduct their every day duties. Employees are able to travel or relocate to any other Forest Service location and never need to learn a new system. The system enables the Forest Service to be more productive and share information faster, all at reduced costs to the public. Independent consulting firms have estimated a benefit/cost ratio of five to one.

3. Typesetting Needs

Quality documents are an integral part of the Forest Service mission. These documents include reports to the public, the Congress and other oversight agencies. The Forest Service needs to produce the highest quality printed documents. The research community has needed a system that is available to everyone that can produce manuscripts with Greek letters and mathematical equations. Our desire is to use our existing system and not implement another technology strictly to fulfill this need. \TeX on the Data General provides this capability. Integration into the CEO system has greatly simplified the use of \TeX . Forest Service employees can use \TeX to produce high quality printed documents knowing little or nothing about \TeX itself. CEO commands are translated into \TeX while \TeX commands are simply passed through as the document is exported to \TeX . The conversion and export processes are transparent to the user.

The integration of \TeX into our existing word processing system meets our goal of providing a consistent user interface and does not require all employees to learn a new system.

4. \TeX Interface

\TeX is used by inserting commands into a document. These commands instruct the \TeX system how to produce the desired printed document. The results of the commands are not seen until the document is printed. The CEO word processing system, on the other hand, is almost a WYSIWYG (What You See Is What You Get) system: it performs functions by displaying the text in the selected format at the time the function is entered. For example, a line centered in CEO is displayed with the center symbol in column one and the correct number of spaces to make the text centered. In short, these systems use two entirely different methods for processing textual information.

`CEO.to-TEX` is a program developed to convert a CEO document into a file to be processed by \TeX . This program will convert many of the CEO functions (e.g., center, text attributes, indents, etc.) as closely as possible to equivalent or comparable \TeX commands. It is important to note that the \TeX system, unlike the CEO system, uses proportional spacing to produce high-quality documents. The conversion program gives the user as much control as possible via CEO functions without degrading the quality of the \TeX output. By doing this, the user can produce high-quality typeset documents from CEO without having to know any of the \TeX commands. It can also be used by knowledgeable \TeX users to convert existing CEO documents into \TeX format without having to re-type the entire document.

For some users, the conversion program may be all that is needed to produce a final product. For others, it may be just the first step in the process.

4.1 Features of the `CEO.to-TEX` Conversion Program

For many of the CEO features, the conversion to \TeX is a straightforward process. However, there are certain conversions that require special consideration. The following sections describe these special considerations.

Special T_EX symbols

The T_EX system uses the following infrequently used typewriter symbols in its command syntax:

\ { } \$ & # ~ - % ~

The \, { and } symbols are the most frequently used of these for basic T_EX commands and infrequently used in most word processing documents. Thus, we made the decision to sacrifice these symbols so that the CEO user could more easily embed T_EX commands in a document. The remaining symbols are converted to the equivalent T_EX symbol. The system was designed with the flexibility to allow the user to convert or not convert any of these symbols. By default, the \, { and } are not converted; all others are converted. The fact that the \, { and } are not converted allows the CEO user to embed T_EX commands directly into a document. For example, the following will result in the text between the brackets printing in a monospaced typewriter text font:

The use of {\tt monospaced fonts} can be quite useful for software documentation. This example illustrates the fact that the three symbols are passed through to T_EX with no conversion to represent the appropriate T_EX commands. Any of the special symbols can be converted or not converted independently of one other. To do this, the user simply includes `symbol/ON` or `symbol/OFF` in a CEO user note prior to using the symbol. For example, the following sequence illustrates how to use T_EX math mode in a CEO document:

```
!$/OFF!  
$$ Y = \alpha + \beta X $$  
!$/ON!
```

The exclamation points (!) represent a CEO user note. The user can turn any symbol conversion off and on anywhere in the document. If a user wants to use the \, { or } symbols in a document, the `symbol/ON` user note can be used to allow the symbol to be converted and not represent a T_EX command.

CEO tabs and indents

Tabs should only be used to produce tables, and as a paragraph indent. When tab stops are used, the most recent format ruler (current) must specify the tab stops required and only those stops. They must be used correctly in conjunction with the format ruler. Every tab stop in a line must be used when typing text even if the column is blank. A common error made when inserting tabs or indents is to type over existing tabs in one line and then tab past them in another line. This is a definite mistake when using the conversion program because it uses the current format ruler to construct T_EX-compatible tab stops.

Tabs and indents were designed for specific functions and should not be used interchangeably. Tabs are used to position information in columns of a table and to start the beginning of a paragraph. Indents should never be used in a table. Indents are used to offset information from the left margin. All information following an indent will be offset to the current tab stop until the end-of-line character is encountered. Tabs and indents should never occur in the same line and/or paragraph of text.

Blank lines in tables

Whenever the conversion routine encounters a tab character in any column other than the paragraph indent column (column 5), it switches to table mode. If a table is to contain blank lines, they should be entered with all tab stops specified. If the blank lines between lines of text in a table do not have tab stops in them, T_EX will switch in and out of table mode for every blank line of the table. It is more efficient to specify null tabs in blank lines and maintain table mode than to switch T_EX in and out of table mode.

Spaces in tables

The T_EX system normally compresses all spaces to one space and removes leading spaces in a table. To allow some flexibility, the conversion routine converts all spaces in tables into hard spaces. Thus, unlike T_EX, CEO spaces can be used to align information within columns of a table.

CEO hard space

The CEO **hard space** feature can be used to make sure words with spaces in them always print on the same line. For example, the words "Dr. Childs" should always occur on the same line of print. To ensure that this always happens, insert the CEO hard space character between the two words. The CEO hard space will be translated to \TeX and will perform the same function as in CEO.

CEO block protect

In CEO a block of text can be marked with the **block protect** feature. If the block of marked text cannot be printed on the current page, it will be started on the beginning of a new page. This feature is useful to ensure that a complete table or other block of text is printed on the same page (provided the entire table fits on one printed page). **block protect** can also be used to eliminate orphaned headings (paragraph or section heading located at the bottom of a printed page) by marking the heading and one or more lines of the section body. You should be careful not to mark large blocks of text with **block protect** because it may cause unnecessary white space to occur on the previous page.

CEO end page

CEO has two end-of-page types; soft and hard. A soft end-of-page is represented by a dotted line across the screen and automatically occurs every time the number of lines specified by the current format ruler is encountered. A hard end-of-page is represented by a solid line across the screen and is manually inserted by the **end page** function key. Soft end-of-pages are not converted to \TeX , as \TeX will automatically page text based on the current page size specification. Hard end-of-pages will be converted to \TeX page ejects.

CEO include

The CEO **include** function is used to attach another document to the one currently being processed. The document to be included does not get inserted into the current document at the time the **include** function is invoked. Instead, the path name of the document is displayed inside the **include** symbols. At print time, the contents of the **include** document are inserted into the current document at the point where the **include** occurs. This feature is processed in the same manner by the \TeX conversion routine.

include documents can contain other **include** documents; this is called nested includes. For example, a paper from the proceedings in the above example might contain five chapters and each chapter might be a separate document specified by an **include**. So the main document contains **include** documents of which one or more contains **include** documents. The conversion routine was tested to 200 levels of nested includes before it was terminated.

CEO footnotes

The CEO system provides an option for specifying footnotes in a document. The conversion routine will extract the footnote information from CEO and convert it to the \TeX system.

CEO headers and footers

The CEO page headers and footers are used to put additional information, such as titles and page numbers, at the top and/or bottom of a printed page. CEO allows you to put up to six lines of text in both the headers and footers; however, it does not allow the use of text attributes, such as bold, underline, etc., but does allow text to be tabbed and centered. Also, the user can specify any valid \TeX command in a CEO header and footer, which will be transferred directly to the \TeX headline and footline.

CEO bold and underline

CEO provides a feature for bold, underline, bold underline, double underline and bold double underline. Underlining, however, is generally not used in typesetting because it does not look good:

For example, underlining dosen't always look very good.

The underline either smashes descenders or is spaced too far from words without descenders to compensate for words with descenders. The convention we have used is to convert underlined text to

italicized text and double underlined text to slanted text. The conversion program uses this convention for underlined and double underlined text. In addition, bold underlined text is converted to bold italicized text and bold double underlined text is converted to bold slanted text. Bold text in CEO is converted to bold text in T_EX.

CEO center

The CEO `center` function is converted to a T_EX `\centerline`, which produces the same results.

CEO quote marks

In CEO, the double quotation is used for both the left and right double quote marks, whereas in typesetting, there is a difference between the left and right quote marks: the left double quote marks look like two backwards (and inverted) apostrophes (“) and right double quote marks look like two apostrophes (”). The conversion routine will keep track of CEO double quotes and convert them into the appropriate left or right quotes. The conversion routine assumes a double quotation immediately following a digit is an indicator for inches and does not convert it to either a left or right quote. If in fact, the double quote following a digit should be either left or right quote marks, it can be specified by turning T_EX mode on and entering either ‘‘ for right double quotes or ’’ for left double quotes and then turning T_EX mode off again.

Dashes and hyphens

In typesetting, there is a difference between a hyphen and a dash. In fact there are three different symbols. One is a hyphen, which is used for compound words like X-ray. The second is an en-dash, used for number ranges such as “lines 3–20”. The third is an em-dash (normal dash), which is used for punctuation in sentences — like this. The difference is in the length of the line representing the symbol. The conversion program will pass the hyphen characters directly to T_EX and T_EX will treat one typed hyphen as a hyphen, two hyphens as an en-dash and three hyphens as an em-dash.

CEO mandatory end-of-line

In CEO, the mandatory end-of-line is produced by either the `NEWLINE` or `CR` key. The CEO mandatory end-of-line is used to end a sentence, end a line in a table, end a paragraph, exit indent mode or produce vertical white space in the printed document. The conversion for a mandatory end-of-line will depend on where it is encountered.

- If the mandatory end-of-line is encountered at the end of paragraph, the conversion routine will instruct T_EX to terminate the current paragraph, produce a specified amount of white space and begin a new paragraph. The new paragraph will be indented if the next line contains a tab skip to column five; otherwise, the paragraph will not be indented.
- If a mandatory end-of-line is encountered while in table mode, the current row of the table will be terminated.
- If a mandatory end-of-line is encountered while in indent mode, the conversion routine will instruct T_EX to exit indent mode.
- When two mandatory end-of-lines are encountered in sequence, the conversion routine will instruct T_EX to produce white space equivalent in vertical dimension to a line of text.

Superscripts and subscripts

CEO superscripts and subscripts will be converted to T_EX equivalents.

Overstrike mode

CEO supports an overstrike mode that allow characters to be printed on top of other characters. The conversion routine attempts to convert this feature to T_EX; however, it is almost impossible to do because T_EX uses proportional spacing. This feature should not be used. T_EX should be able to produce any character needed.

CEO discretionary hyphen

CEO supports discretionary hyphens, which tell the software where a potential line break in a word

can be made. The conversion routine translates CEO discretionary hyphens to T_EX discretionary hyphens, which perform the same function as in CEO. If a line break does not occur in the word, the discretionary hyphen is ignored.

4.2 Additional Conversion Features

The conversion program contains additional features which allow CEO users greater flexibility in producing typeset documents.

Special symbols

The following table contains a list of special symbols that are not part of the CEO character set but can be produced in T_EX documents. These symbols are specified with the CEO superscript text attribute. For example, the bullet symbol can be produced with a superscripted period, i.e., ^ˆ.^ˆ. Note that the ^ˆ symbols enclosing the period (.) are produced by the superscript text attribute and not by the keyboard character.

Superscript text	Description	CEO Examples
^ˆ . ^ˆ	Bullet	• This line starts with a bullet.
^ˆ R ^ˆ	Registration symbol	This is a CEO® example.
^ˆ C ^ˆ	Copyright	Copyright © 1988 by Data General Corp.
^ˆ T ^ˆ	Trademark	The PRESENT TM utility is a versatile tool.
^ˆ o ^ˆ	Degree Symbol	The temperature was -25°F today. (This is a lowercase o, not a zero)
^ˆ 1/2 ^ˆ	Fraction	It rained 1½ inches today.

Fractions can be produced using the superscript text attribute and any digits separated by a slash, for example, ^ˆ1/3, ^ˆ5/16, ^ˆ19/64, ^ˆ3/323, etc.

Special spacing features

When a period is encountered by T_EX, it is assumed to be the end of a sentence and a space representing end-of-sentence spacing is put after the period. However, if the period is not the end of a sentence, as in "Dr. Childs", then the amount of space inserted is out of proportion with the rest of the spacing in the sentence. The conversion program scans the CEO text for the character strings Dr., Drs., Ph.D., Mr., Mrs. and Ms., and inserts a single space instead of end-of-sentence spacing.

5. ASCII Previewer

We have also provided an ASCII previewer to get a crude representation of what the typeset document will look like on our existing ASCII terminals. This screen representation allows the user to see page breaks, line endings and orphaned text. The previewer software was designed following the CEO user interface with consistent screens and function key mappings. Figures 1 and 2 demonstrate the user interface developed for implementation of T_EX in the Forest Service.

6. Data Tables

CEO decision support tools include a data table product that allows a user to enter data in a table format. In keeping with our desire to use existing products, a conversion of CEO Data Tables to T_EX-formatted tables is provided. This conversion takes data from the data table and builds the appropriate T_EX commands to produce a table. The user can select options to include table captions (above or below), column headings, and horizontal or vertical rules.

```

Msg: New: 0                               Mar 28,89 5:26 PM           username:hostname
                                           FOREST SERVICE TEX

      ---> 1. Typeset      (Typeset and print CEO document)
           2. Preview      (Display typeset document)
           3. View          (Display CEO document)
           4. Edit          (Revise CEO document)
           5. File          (Create AOS/VS file)

Enter Choice:1

To return to the previous menu, press the CANCEL/EXIT key.
For assistance here (or on any other menu or question), press the HELP key.

```

Figure 1: T_EX environment main menu

```

Msg: New: 0                               Mar 28,89 5:26 PM           username:hostname
                                           FOREST SERVICE TEX

Document: TeX test                          in
Drawer: TeX drawer                          Folder: TeX folder

Pick one: (1. TeX, 2. LaTeX, 3. SliTeX, 4. BibTeX) 1

Convert to TeX format (Y/N)? Y

Printer name:

How many copies? 1

Run in background (Y/N)? Y

Execute (Y/N)? Y

```

Figure 2: T_EX environment typeset menu

7. Using Graphics

Graphics produced by either CEO Drawing Board, Charting Tool, Present or application programs using the Graphics Kernal System (GKS), in either GKM or CGM format, can be inserted into a T_EX document. The procedure for including a graphic consists of two parts; defining the graphic to T_EX and preparing the graphic for inclusion.

8. Documentation

Documentation describing the Forest Service implementation of T_EX is available from the authors. It includes a description of T_EX, the Forest Service environment, CEO to T_EX conversion features, previewing text on ASCII terminals, building tables with CEO Data Table, using graphics in a T_EX document, and a description of other T_EX products. Also included are appendices that provide the user with a glossary of technical terms, common T_EX terms, a T_EX font chart, Forest Service T_EX macro library, a detailed description of using T_EX in the Forest Service environment, T_EX defaults for CEO conversion, T_EX software release notice, and T_EX software abstract for the Software Reference Center.

9. Summary

The T_EX and CEO integration software was just released within the Forest Service this past May. Fewer than 50 sites have the software installed, and two national training classes have been held, with more to follow soon. We also expect to release the CEO Data Table and the Graphics Inclusion software by the end of October.

The integration of a sophisticated typesetting system such as T_EX into an existing word processing system produced an extremely powerful document processing tool for the Forest Service. Any employee having access to a Data General system anywhere in the organization can now produce the highest quality typeset documents.

Training consisted of 3 courses similar to the elementary T_EX course offered by TUG. Others will be offered as our user community grows. Distribution will be in the form of magnetic tape, and also electronically through the Forest Service network. Distribution will include the T_EX executables, the CEO_to_TEX conversion program, several Forest Service-specific macros, and the Computer Modern and other fonts. Additional macro packages will be distributed as the need is identified.

Integration of these two systems has allowed the Forest Service to maintain the integrity of its information processing standards developed over many years. This is viewed as a major accomplishment in an organization the size and complexity of the USDA—Forest Service. We hope to make a future report on the use of T_EX in the Forest Service.

10. Acknowledgements

The authors would like to express their sincere appreciation to Dr. Bart Childs, Texas A&M University, for making T_EX available on the Data General systems and for his continued support throughout this project; and to Dr. Donald Knuth, Stanford University, for developing T_EX and putting it in the public domain so projects like this could become a reality.