MONASH UNIVERSITY

MULTILINGUAL AND MATHEMATICAL TEXT PROCESSING

Esme Preston

DEPARTMENT OF ECONOMETRICS

MULTILINGUAL AND MATHEMATICAL TEXT PROCESSING

Esme Preston

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS AND POLITICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

MULTILINGUAL AND MATHEMATICAL TEXT PROCESSING

BY

Esme Preston

Developments in multilingual and mathematical text processing are proceeding apace; it is not easy to keep up with them. It is essential, both to purchasers and to software developers, that there should be vigorous discussion of this major area of current software development. The reviews of OK Editor by Ralph [1] and Cranswick [2] have brought to my notice what is clearly a product of merit. This paper is intended as a contribution to further discussion of technical wordprocessing software and a demonstation or description of some of the alternative products.

The OK Editor output of foreign languages displayed by Ralph (p. 108) and further material supplied by the vendor on request is of excellent standard. According to Cranswick [2], the original program was written by a physicist, European characters were 'designed by the University of NSW' and a facility for overstriking characters, needed for some mathematical text, 'comes from Macquarie University'. With such distinguished origins it is not surprising that the OK Editor reaches such a high standard of output.

In the area of mathematical text and the special structures required by scientists, however, the display provided by Ralph, though of good standard, is by no means sufficient to allow an

1

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

assessment of the relative merits of OK Editor vis-à-vis its competitors. There are internationally established standards, described below, which must be met by technical wordprocessors. Neither Ralph nor Cranswick appears to be aware of these standards or of the extensive literature created to serve both purchasers and developers of this type of software. OK Editor may well meet these standards but Ralph's review does not enable the reader to make a judgement. Furthermore, Ralph's closing comments are sadly wide of the mark; to suppose that the Australian product is 'without serious competition' or that purchasers 'have no real choice' is patently absurd. There *are* other excellent products available for processing foreign language or mathematical text. The question is how well does the Australian product stand up to fierce international competition in this area. Our TV's are constantly urging us consider Australian products but not, as yet at least, to bury our heads under Ayer's Rock.

This article discusses some of the serious competition and alternative choices available; it also displays some of the standard benchmark tests which potential purchasers need to apply to software products designed for the preparation of mathematical text. The article does not attempt to assess, except incidentally, OK Editor's capabilities already reviewed by Ralph and Cranswick. The same product used to prepare the benchmark display provides facilities for multilingual text-processing which we shall also demonstrate. Neither Australian purchasers nor Australian software producers can afford to be ignorant of such

2

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

strong international competition.

The internationally accepted standards for technical word processors (TWP) to which we refer are those established by the American Mathematical Society and the Boston Computing Society beginning early in 1986 and elaborated in 1987; these are fully discussed in [3] and [4]. The *Notices of the American Mathematical Society* has had, for some time, a regular column devoted to Technical Wordprocessing under the editorship of Richard S. Palais. Contributions to this column come from distinguished scientific writers all over the world. Individual reviewers of TWP are hard put to it to compete with this authoritative source of information. Indeed, software reviewers in some of the best-known personal computing publications have been strongly criticised in this column. Thus, for example, *Lotus Manuscript* is one contender discussed by, [5], [6], [7] and [8]. The Boston Computer Society comments on *Manuscript* and its reviewers. Of the product the BCS says 'not ready yet, even though this is what they are selling'; and of its reviewers: 'It is obvious that neither Machrone nor Seymour has any real understanding of what a TWP is' while Stewart's is 'not a good preview'. See p. 490 of [9].

The standards set by AMS-BCS currently include every aspect of technical manuscript preparation but began with (and still retain) 10 equation-type benchmarks fully displayed in [3]. Two of these, prepared by means of the TWP ChiWriter, are reproduced below. Few TWP are able to meet such exacting standards;

3

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

comparative tables showing results for these and many other tests applied to more than twenty such products are provided in [10]. One product capable of satisfying all these tests is T_EX. The American Mathematical Society currently (beginning 1987) invites contributors to submit papers in the T_EX file format. Since the Australian Mathematical Society is about to adopt T_EX for its publications as of 1988, it may be confidently expected to follow suit. Other academic journals will surely do likewise in the near future.

Mainframe and PC versions of T_EX are currently in use at Monash. Other products used at Monash include T^3 and ChiWriter. This paper discusses the last of these in some detail. ChiWriter satisfies the 10 benchmark tests mentioned above (as does T^3) though it does not have the typesetting quality or manuscript-managing facilities of T_EX.

The idea that mathematicians are particularly interested in wordprocessing seems to strike many people as strange yet the connexion between mathematics and the preparation of fine documents goes back many centuries. In the development of electronic typesetting and wordprocessing the 'seminal' article is that by the mathematician Knuth [11] who designed T_EX for typesetting technical material. Knuth's paper constitutes an introduction to typesetting excellence and the mathematics needed for electronic typesetting; it has a bibliography of references reaching back to 'about 1460'. Knuth's books [12] and [13] (The $T_EX$book and The METAFONTbook) are internationally regarded as

4

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

definitive. T_EX is available for the IBM PC and clones in versions known as PC T_EX and MicroT$_E$X. Both have agents in Australia[1]. Macro packages which make T_EX easier to implement are fully described by Spivak [14] and [15], and Lamport [16]. A first introduction to T_EX is provided by [17]. All these authors use a light-hearted style which is a welcome relief from the turgid style of most wordprocessing manuals and indeed most computing manuals.

For some users the markup language used by T_EX is a problem. Many people think visually rather than verbally. Odd as it may seem to non-mathematicians, Hadamard in his book [*The Psychology of Mathematical Invention* (1945), Jacques] reports of mathematicians and scientists that their mental pictures of mathematics 'are most frequently visual'. A brief account of Hadamard's findings can be found in [18]. This visualising certainly does not mean seeing algebraic symbols as it might be on a display screen but something much vaguer. Could it be, however, that the visual approach to technical wordprocessing is more closely related to the mathematical way of thinking than the markup language? From experience it would seem so. Those who, for whatever reason, prefer a WYSIWYG (What You See Is What You Get) showing output immediately to the screen as the document is prepared are well catered for by ChiWriter which has considerable

---

[1]   PC T_EX:The Wordworks, The Boulevard Lawns, City Walk, ACT,2601. MicroT$_E$X:Addison Wesley, Sydney.

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

power for multilingual and mathematical wordprocessing as we shall show below.

OK Editor has a number of features not offered by ChiWriter or by any of the other top-flight technical word-processors we know of. The latter, for example, has no built-in communications package such as is provided with OK Editor. (The Public Domain program KERMIT is used locally to transfer files.) Nor does ChiWriter offer built-in numerical functions corresponding to those provided by OK Editor; it can be supplemented by a spreadsheet capable of extracting to disc. One particularly attractive feature of OK Editor is its ability to control horizontal half-spaces which gives a much more pleasing appearance to many equations; ChiWriter is unable to do this though other well-known TWPs offer this feature. A single copy of ChiWriter running in CGA mode costs $US79.95[2] which makes it very competitive.

Obversely, ChiWriter has features not offered by OK Editor. In particular, it is menu driven. This feature is illustrated on the following page produced by a simple screen dump to an Epson LX-80. The menu commands shown at the bottom of the screen, are accessed either by the [Esc] and arrow keys or by the [Alt] key combined with capital letter. The menu is tree-structured,

[2] Chiwriter is supplied by Horstmann Software Design Corporation, 140 E. San Carlos Suite 200, P.O. Box 5039, San Jose, CA 95150, USA. A Demo Disk is available.

as having 'standard Roman, foreign (all western European and

Scandinavian languages), classical-koine-mode Greek, and

biblical-modern Hebrew alphabets'. For French or German typists a

problem arises in the layout of the keyboard (French, for example,

is AZERTY not QUERTY). There are now available German[9] and

French[9] versions of ChiWriter with keyboards, manual and commands

all adapted to the language and its typing conventions.

→ChiWriter fonts are called by the function keys F1, ..., F10,

branching down from the main menu - the branch from Screen, for example, is shown in the second line of the menu when the Screen command is highlighted. Mouse support is available for those who wish it. The top line of the illustrated screen shows the filename, the active font, the proportion of memory in use and the modes (synchronised, insert, justified, double spacing, row, column, and page of the cursor location). A menu mode of operation together with the ample help files and on-disc tutorial render ChiWriter particularly easy to learn. The latest update (Version 2.52) which has just arrived on my desk also has notepad and graphical import features which promise to be useful. This version has grown a little beyond the single-disc size of the original system. An outstanding feature of ChiWriter is the close relation between its saved-to-disc files and those needed for T_EX but this goes beyond the scope of the present discussion.

The standard version of ChiWriter comes equipped with a 'foreign' font with numerous non-English characters and also has a Greek font. The output of each is displayed below. It also has a font designer which is not difficult to use and allows the user to add to the set of characters. Around the world, scientists and linguists have contributed to the development of further fonts which are offered as shareware - see below. One Cyrillic font has been designed by a Melbourne 13-year-old and another Melbourne user plans to design an Ukrainian font. The vendor-supplied foreign characters appear to include most of those required for Ralph's Vietnamese example and key sequences (up to 50 can be

accomodated) or font designer could supply the rest.

Until recently ChiWriter has not been advertised and has become known by personal recommendation and through User Group Bulletin Boards. The current special issue of BYTE [19], however, carries an advertisement (p.234) which provides a price list for various vendor-supplied add-ons; the same issue also has a short account (p.11) of 'The Scholar's Edition' of ChiWriter described as having 'standard Roman, foreign (all western European and Scandinavian languages), classical-koine-mode Greek, and biblical-modern Hebrew alphabets'. For French or German typists a problem arises in the layout of the keyboard (French, for example, is AZERTY not QWERTY). There are now available German[3] and French[4] versions of ChiWriter with keyboards, manual and commands all adapted to the language and its typing conventions.

ChiWriter fonts are called by the function keys F1, ..., F10, or Shift and function key 'chord' for higher numbered fonts as indicated in the table below. For each of the fonts displayed the lower case is shown first followed in the next row by the upper case. The table has been arranged to correspond with keyboard locations rather than alphabetically so that the reader, by comparing characters with the standard set, may see something of

[3] Distributed by Hofbauer Informationssysteme, Neunkircher Str. 18-20, A-2700 Wiener Neustadt, Austria.

[4] Distributed by Churing, S.A.R.L., Rue de Seine, 75006 Paris, France.

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

the way in which characters are allocated to keys. A careful study of the data will show that many of these allocations are highly mnemonic. Consider, for example, which qwerty (Standard font) keys correspond with ∞ ∃ ∀ ⊂ ⩾. For the learning stage, the screen representations of the keyboard allocations of the various fonts (called by Alt-H after a function key) are very helpful.

The table on the following page shows all the characters supplied in the 'as is' version of the ChiWriter program as printed by the HP Laser Jet Series II printer. Other less expensive printers are supported including 9-pin Epson LX86 and the Toshiba 24-pin dot matrix printers. There are vendor supplied key-sequences for various large symbols such as those at the bottom of the font display and a utility for the user to construct others. Such 'macros' can be used, for example, to combine the items provided by MathII font or to construct accented letters required for non-English languages. Many other fonts are available from users; for the most part the developers are 'academics'. Horstmann offers to distribute as shareware suitable fonts developed by users.

Chiwriter Fonts Printed on HP Laser Jet Series II

F1 Standard `'1234567890-=\qwertyuiop[]asdfghjkl;'zxcvbnm,./`

~!@#$%^&*()_+|QWERTYUIOP{}ASDFGHJKL:"ZXCVBNM<>?

F5 Foreign ` àøæ þ°¯ šëéêèüúûùöóôòäáâàïíîìℓđ· ‡ç·ßñõãÿ

¯¡ÅØÆ · Þ ŠËÉÊÈÜÚÛÙÖÓÔÒÄÁÂÀÏÍÎÌℒÐ· ‡Ç· ÑÕÃŸ¿

F6 Symbol "*†‡ —   ƒ   " ¢   O

£   ®™¥ ¶ §   © ▷

F7 Greek θωερτϑυιοπ ασδφγηψκλ ζξχφβνμ

ΘΩΕΡΤ ΤΙΟΠ ΑΣΔΦΓΗΨΚΛ ΖΞΧ ΒΝΜ

F8 Linedraw

F9 MathI

F10 MathII QWERTYUIOP{ }ASDFGHJKL ZXCVBNM

F2 Small qwertyQWERTY   F3 Italic *qwertyQWERTY*   F4 Bold **qwertyQWERTY**

F11 Underline <u>qwertyQWERTY</u>   F12 Orator QWERTYQWERTY

F13 Script *qwertyQWERTY*   F14 Gothic qwertyQWERTY

Among potential users of software there seems to be a general suspicion of low-priced software which is often assumed to be the handiwork of 'hobbyists' rather than professionals. It is true that there are traps for the unwary; on the other hand the international user network offers many excellent items of software which are either free or very inexpensive. Many of them are provided or offered for sale by internationally known academic staff in various areas of expertise; others by computing experts in the non-academic community. ChiWriter is one of the former. The author is a well known professional mathematician and many other computing experts and distinguished academics around the world are taking a hand in ChiWriter's development. Users active in developing fonts or providing other facilities for extending ChiWriter include mathematicians, scientists, philosophers and economists in the following institutions: Department of Economics, Queen's University, Canada; Computer Science Department, Washington University, USA; Department of Physics, Northwestern University, USA; University of Heidelberg, West Germany; Department of Physics, University of Maryland, USA; University of Geneva, Switzerland; Bedford Institute of Oceanography, Canada; University of Paris VII, France; Department of Philosophy, La Trobe, Australia. From recent developments it is clear that linguists are also becoming interested in the linguistic capabilities of ChiWriter and active in developing them. See [19].

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

Two of the 10 equation-type AMS-BCS benchmarks are displayed in the table below. The tenth benchmark is the inclusion of all the other nine in a single document so the present document demonstrably goes some way towards meeting this also. For the remaining tests of the same sort see [3].

An independent review of ChiWriter is provided by [20], while [21] has some further comments on $T_E X$. The $T_E X$ User Group has its own newsletter TUGboat. The ChiWriter Chronicles is the name of a rather infrequent newsletter provided by Horstmann. Some further remarks on Technical Wordprocessors my be found in [22].

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

Two Benchmarks of the American Mathematical Society

and the Boston Computer Society

Prepared Using ChiWriter

Benchmark 6

Allen F. Henry, *Nuclear Reactor Analysis*, MIT Press, Cambridge, Mass, 1982, page 495

$$iB_r[\tilde{a}_{k1}^n] \equiv \frac{1}{2}(h_{n-1} + h_n)\int_0^R 2\pi r dr[\rho_k^{n*}(r)]\frac{d}{dr} [\Psi_1^n(r)],$$

$$[D_{r,k1}^n]^{-1} \equiv \int_0^R 2\pi r dr \int_{z_n -\frac{1}{2}h_n -1}^{z_n +\frac{1}{2}h_n} [\rho_k^{n*}(r)][D^{-1}(r,z)][\rho_1^n(r)],$$

Benchmark 7

Richard P. Feynman, *The Feynman Lectures in Physics*, Addison-Wesley Publishing Co., 1965, Vol 3, page 20-12, Table 20-1.

Physical Quantity	Operator	Coordinate Form
Energy	$\hat{H}$	$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m}\nabla^2 + V(r)$
Momentum	$\hat{p}_x$	$\hat{\mathcal{P}}_x = \frac{\hbar}{i}\frac{\partial}{\partial x}$
	$\hat{p}_y$	$\hat{\mathcal{P}}_y = \frac{\hbar}{i}\frac{\partial}{\partial y}$
	$\hat{p}_z$	$\hat{\mathcal{P}}_z = \frac{\hbar}{i}\frac{\partial}{\partial z}$

14

REFERENCES

[1] Jeff Ralph, OK - More Than Just 'All Right', *Australian Computing*, Summer '88 Edition, 104–108.

[2] Kester Cranswick, *OK Word Processor*, Australian Personal Computer, July 1987, 1–5.

[3] Boston Computer Society PC Technical SIG, "Technical Wordprocessors for the IBM PC and Compatibles", *Notices of the American Mathematical Society*, 33 (1986), 8 - 37.

[4] Richard Goldstein, James Loomis, and Avram Tetewsky, "Technical Wordprocessors for the IBM PC and Compatibles, Report by the Boston Computer Society: Part I - TWP Capabilities and People Needs", *Notices of the American Mathematical Society*, 34 (1987), 15–32.

[5] B. Machrone, *Manuscript Whips Large Documents into Shape*, PC Magazine, February 20, 1987, 33–35.

[6] J. Seymour, *Manuscript: Something Old, Something New*, February 20, 1987, 85–87.

[7] G. A. Stewart, *Lotus Manuscript*, BYTE, November 1986, 91–101.

[8] E. Bender, *Jonathan Sachs After 1-2-3*, PC World, December 1986, 212–219.

[9] Richard Goldstein, James Loomis, and Avram Tetewsky, "Technical Wordprocessors for the IBM PC and Compatibles, Report by the

E.Preston,D E O R,Faculty of Economics,Monash,Vic 3168

Boston Computer Society: Part IIB - Reviews", *Notices of the American Mathematical Society*, 34 (1987), 462 - 490.

[10] Richard Goldstein, James Loomis, and Avram Tetewsky, "Technical Wordprocessors for the IBM PC and Compatibles, Report by the Boston Computer Society: Part IIA - TWP Summary Tables", *Notices of the American Mathematical Society*, 34 (1987), 262 - 281.

[11] Donald E. Knuth, "Mathematical Typography", *Bulletin (New Series) of the American Mathematical Society*, Volume 1, Number 2, March 1979, 337-372.

[12] Donald E. Knuth, *The $T_EX$book*, Addison Wesley and the American Mathematical Society, 1984, Sixth Printing 1986.

[13] Donald E. Knuth, The METAFONTbook, Addison-Wesley and the American Mathematical Society, 1986.

[14] M. D. Spivak, *The Joy of T_EX, A Gourmet Guide to Typesetting with the AMS-T_EX macropackage*, American Mathematical Society, 1986.

[15] M. D. Spivak, *PC T_EX Manual*, Second edition, PC T_EX Inc, 1985.

[16] Leslie Lamport, *L^AT_EX User's Guide and Reference Manual*, Addison Wesley, 1986.

[17] Arthur L. Samuel, *First Grade T_EX - a Beginner's T_EX Manual*, Report No STAN-CS-83-985, Stanford University, Department of Computer Science.

[18] Philip J. Davis and Reuben Hersh : *The Mathematical Experience*, Penguin, 1983.

[19] BYTE Extra Edition, *Inside the IBM PCs*, 12, no. 12, 1987.

[20] G. Cornell, *Review of ChiWriter*, The Mathematical Intelligencer, 8, no. 2, 64.

[21] William Abikoff, T_EX - *Six Months Later*, The Mathematical Intelligencer, 9, no. 1, 66-67.

[22] Esme Preston, *On Choosing a Technical Wordprocessor*, The Australian Mathematical Society Gazette. To appear.

Esme Preston

Dept. of Econometrics and Operations Research

Faculty of Economics

Monash University

Clayton, Vic., 3168