

BIJLAGE I**A dBase III+ programme to generate a journal****Piet Roes****International Statistical Institute****Pr. Beatrixlaan 428****2273 XZ Voorburg**

This is a note about the implementation of a combination of dBase III+ and T_EX. From the 1989 volume onwards the *Statistics Theory and Methods Abstracts (STMA)* journal is being typeset by its editors. For this purpose the permanent office of the *International Statistical Institute* at Voorburg has purchased an Olivetti 380 S with a 40 Mb disc, a tapestreamer and a HP laserprinter.

The journal production process, has been streamlined considerably and now proceeds roughly as follows. Abstracts on relevant papers are written and classified as soon as new journals arrive. The classification is in accordance with a scheme of some 260 classes. Every three months approximately one thousand abstracts are put out. The papers abstracted have one to four authors, an address of the first author, a title, a language, journal information: year, volume, issue, first page, last page and the number of references. Together with this information, the abstract is included in one record. The abstract varies in length from a few hundred to more than one kilobyte. Naturally, the abstract is put in a "memo" field. Title as well as abstract may contain any sort of maths.

The maths mentioned at the end of the last section persuaded us to go for AMSTEX. The use of a database is dictated by the various ways a reference journal must provide entries to its subject matter. The abstracts are presented in the order of their classification number. This has to be arranged at the very last minute before printing, because material becomes available in random order. For a sequence of abstracts to be satisfactory, there must be sorting on journal, year, issue, first author and first page as well.

For the right emphasis in context, authors names are displayed in capital letters, but occur in the author index in lower and upper case. This causes some problems, as many foreign names with foreign accents and foreign conventions have to be dealt with. Also, authors names have to be listed alphabetically from time to time. This aspect is implemented by deriving a capitalised name through standard conversion except for T_EX codes like `\b \c \d \u \v \t`. For the sake of ordering the authors alphabetically, their names are stripped of all nonletters and are capitalised subsequently. Further, words like Mac and jr. are not capitalised. To this end the latter are to be input as `M%a%c` and `%j%r`.

All regularly recurring journal names, addresses and text are coded for efficiency. Thus there is a list of affiliations and a journals list. Incidentally, the latter is kept also to be output yearly and for keeping track of their occurrence through the year. Also, full use is made of the coding facilities which the editor (PC write) provides, while the remainder of the standard words like "probability" or "estimator" are coded using T_EX.

There are several ways the output of the database can be sent to T_EX. For the sake of proof corrections it must be in order of input, either all abstracts or only those actually corrected. The final output for an issue is about 1.3 Mb. That is fine if impeccable, but in order to enable any final correction, it is broken up into six files. We have not yet found an editor that can handle files beyond memory capacity, except View on the old Acorn BBC.

Another problem met with DOS concerns switching from dBase to T_EX. It appears that the

RAM-memory often was not fully released for use to $\text{T}_{\text{E}}\text{X}$ and memory shortage was encountered. This has been resolved satisfactorily by the timely release of a set of software called TSR Utilities Version 2.5. The TSR Utilities include programs useful in managing DOS memory, and in particular managing memory-resident utilities. TSR stands for "Terminate and Stay Resident". Before entering dBase a mark is set in memory; after leaving dBase, the memory above the mark is released.

The $\text{T}_{\text{E}}\text{X}$ source file is actually produced by dBase in a fairly roundabout way. There appears to be no straightforward method to send dBase output to a file rather than to the printer if the output is anything less than standard. Use is made of a sort of log file called "alternate". This is initiated by issuing "set alternate to ABSTRACT.TEX", "set alternate on". The text subsequently scrolls over the screen, all 1.3 Mb of it! There appears to be no way to avoid this.

Apart from the previous minor inconvenience, several problems of similar type were encountered with dBase. One which initially caused great problems was an unlisted anomaly with "set order to order 0", This is supposed to make the database available in input order, while keeping current index files uptodate. The latter does not materialize. Ashton-Tate has not been forthcoming with a way around this; we use "reindex" at present.

A properly documented inconvenience is the illogical way a macro is dealt with by the "do while" command. A macro is a substitution convention: if I=3, AUTHOR&I is AUTHOR3. A "do while" loop however does evaluate such a macro only once.

Finally, there have been some odd records that could not be persuaded to be included in an index for no apparent reason. The latter is on the whole a rather atypical occurrence with dBase III+. It is very reliable and includes many safety features to the point of being almost fool proof. Above all, it is extensively used by so many, that available documentation is unsurpassed.