

Steve Porch

Australian Taxation Office Library, Canberra

Scenario

You've just found the perfect case/article/report on an online service and captured it to disk. You fire up the trusty wordprocessor to print it out, only to discover the file is full of footnotes, unwanted page breaks and other nasties that will (among other things): a) reduce the readability of the document; and b) consume more paper than is necessary. How do you beat this file into shape without spending hours at the keyboard?

With some forethought and planning, it is quite easy to use your wordprocessor's macro facility to automate the more tedious editing tasks in our scenario. Most modern wordprocessing packages have macro capabilities, ranging from simple keystroke recorders to a full programming language. To get the most from such facilities, it pays to spend some time with the manual and experiment to see what can be achieved.

Some tips

- Use keyboard commands rather than a mouse.
- Save each document to its own file rather than capture two or three documents to one large file.
- Work on a backup copy to avoid wayward, untested macros completely trashing a document.
- Manually edit out anything that isn't part of the document before starting. This could include commands that you used to download it, charging information and so on.
- Scroll through the document and make sure each paragraph is separated by a return character.
- Ensure headers near page breaks (such as those used by SCALE and LEXIS) have the same number of lines throughout the document.
- If you wish to turn notes into proper footnotes or endnotes, make a note of the numbering style used: n1 or /1/ are two common methods.

I use WordPerfect 5.1 for DOS so will outline the keystrokes used by this package. With some reading of your wordprocessor's documentation, you should be able to duplicate these macros. Setting up a macro:

To Remove Page Breaks & Headers

- To get rid of unwanted page breaks and their associated headers, move to the very beginning of the document and use Ctrl-F10 to name the macro and turn on the macro recording function.
- Now that keystrokes are being recorded, use Search (F2), then press Ctrl-Enter to tell the search function to find a page break.
- Press F2 again to begin searching.
- When the cursor has moved to the first page break, turn on blocking with Alt-F4 and use the cursor keys to move down over the break and over the header lines.
- Press the Delete key and answer Y at the Delete block prompt, then press Ctrl-F10 to end recording keystrokes.
- To run the macro again, press Alt-F10, enter the name you assigned to this macro and

press Enter. The macro will run and delete the next page break/header combination.

You could keep doing this manually throughout your document, but an easier way is to have the macro repeat itself until it can't find any more page breaks. A quick way of doing this is to use the Repeat function (on my keyboard I have left it assigned to the Escape key). Press Escape (or whatever key you use) then enter a large number, say, 200, and invoke the macro via Ctrl-F10. The macro will then repeat 200 times. A more elegant way is to use (shudder) the Macro Editor to put in the {LABEL} AND {GO} commands that cause the macro to repeat itself.

The completed macro will look something like this:

{DISPLAY OFF}

Prevents macro from displaying

{LABEL}Start~

Sent here by the {GO} command at the end

{Search}{Hpg}{Search}

Searches for page breaks

{Block}{Down}{Down}{Down}

Turns on blocking: blocks break and header

{Del}y

Deletes the marked block

{GO}Start~

Sends execution to the label called Start

Note the three consecutive {Down}s: these are the cursor moves mentioned above and the number will vary according to what you actually enter in your own macro. The macro will repeat until the Search fails to find a page break: at this point your document should now be relatively free of extraneous material. If removing page breaks was all that was necessary you are done.

Footnotes & Endnotes

Footnotes involve quite a few steps. You may have wondered why I recommended ensuring each paragraph be separated by a return character. Given that our document's paragraphs are lines ending in return characters, we can't use commands like *Block to End of Paragraph*, so we need some sort of marker to search for. This is the purpose of said return character: the last line of the paragraph ends in a return, so by searching for two returns together, we know that we have reached the end of a paragraph. One exception to this rule is if a footnote consists of more than one paragraph, don't separate them otherwise this macro will only turn the first paragraph of the note into a footnote. Once you've turned the in-text note into a footnote you will have to edit it later to restore its proper appearance as separate paragraphs.

Other bugbears are that sometimes a note number will appear at the beginning of a line. This requires some minor manual editing to put it in a more appropriate position either before or after running the macro. Notes can appear in the text itself, often separated by lines of dashes from the main text, or in a continuous series at the end of the text proper. Documents divided into sections can also have a separate run of notes at the end of each section. Documents I obtain online often include tables, with notes to the table.

Dealing with all the variants of notes and putting them in their proper place with macros will be the subject of next month's column.
