An environment for multicolumn output*†

Frank Mittelbach
Email: see top of the source file

Printed January 23, 1997

Abstract

This article describes the use and the implementation of the multicols environment. This environment allows switching between one and multicolumn format on the same page. Footnotes are handled correctly (for the most part), but will be placed at the bottom of the page and not under each column. LaTeX's float mechanism, however, is partly disabled in the current implementation. At the moment only page-wide floats (i.e., star-forms) can be used within the scope of the environment.

Preface to version 1.5

This new release contains two major changes: multicols will now support up to 10 columns and two more tuning possibilities have been added to the balancing routine. The balancing routine now checks the badness of the resulting columns and rejects solutions that are larger than a certain treshold.

At the same time multicols has been upgraded to run under LaTeX 2_ε.

I apologise for the state of the code documentation but the work on LaTeX 2_ε kept me too busy to do a proper job. This will hopefully be corrected in the near future.

1 Introduction

Switching between two column and one column layout is possible in LaTeX, but every use of `\twocolumn` or `\onecolumn` starts a new page. Moreover, the last page of two column output isn't balanced and this often results in an empty, or nearly empty, right column. When I started to write macros for doc.sty (see "The doc–Option", *TUGboat* volume 10 #2, pp. 245–273) I thought that it would be nice to place the index on the same page as the bibliography. And balancing the last page would not only look better, it also would save space; provided of course that it is also possible to start the next article on the same page. Rewriting the index environment was comparatively easy, but the next goal, designing an environment which takes care of footnotes, floats etc., was a harder task. It took me a whole weekend[1] to get together the few lines of code below and there is still a good chance that I missed something after all.

Try it and, hopefully, enjoy it; and *please* direct bug reports and suggestions back to Mainz.

*This file has version number v1.5l, last revised 1996/01/13.

†Note: This package is released under terms which affect its use in commercial applications. Please see the details at the top of the source file.

[1] I started with the algorithm given in the TeXbook on page 417. Without this help a weekend would not have been enough.

2 The User Interface

To use the environment one simply says

```
\begin{multicols}{⟨number⟩}
      ⟨multicolumn text⟩
\end{multicols}
```

where ⟨*number*⟩ is the required number of columns and ⟨*multicolumn text*⟩ may contain arbitrary LaTeX commands, except that floats and marginpars are not allowed in the current implementation[2].

As its first action, the multicols environment measures the current page to determine whether there is enough room for some portion of multicolumn output. This is controlled by the ⟨*dimen*⟩ variable `\premulticols` which can be changed by the user with ordinary LaTeX commands. If the space is less than `\premulticols`, a new page is started. Otherwise, a `\vskip` of `\multicolsep` is added.[3]

When the end of the multicols environment is encountered, an analogous mechanism is employed, but now we test whether there is a space larger than `\postmulticols` available. Again we add `\multicolsep` or start a new page.

It is often convenient to spread some text over all columns, just before the multicolumn output, without any page break in between. To achieve this the multicols environment has an optional second argument which can be used for this purpose. For example, the text you are now reading was started with

```
\begin{multicols}{3}
  [\section{The User
            Interface}] ...
```

If such text is unusually long (or short) the value of `\premulticols` might need adjusting to prevent a bad page break. We therefore provide a third argument which can be used to overwrite the default value of `\premulticols` just for this occasion. So if you want to combine some longer single column text with a multicols environment you could write

```
\begin{multicols}{3}
    [\section{Index}
     This index contains ...]
    [6cm]
 ...
```

Separation of columns with vertical rules is achieved by setting the parameter `\columnseprule` to some positive value. In this article a value of .4pt was used.

Since narrow columns tend to need adjustments in interline spacing we also provide a ⟨*skip*⟩ parameter called `\multicolbaselineskip` which is added to the `\baselineskip` parameter inside the multicols environment. Please use this parameter with care or leave it alone; it is intended only for package file designers since even small changes might produce totally unexpected changes to your document.

2.1 Balancing Columns

Besides the previously mentioned parameters, some others are provided to influence the layout of the columns generated.

Paragraphing in TeX is controlled by several parameters. One of the most important is called `\tolerance`: this controls the allowed 'looseness' (i.e. the amount of blank space between words). Its default value is 200 (the LaTeX `\fussy`) which is too small for narrow columns. On the other hand the `\sloppy` declaration (which sets `\tolerance` to $10000 = \infty$) is too large, allowing really bad spacing.[4]

We therefore use a `\multicoltolerance` parameter for the `\tolerance` value inside the multicols environment. Its default value is 9999 which is less than infinity but 'bad' enough for most paragraphs in a multicolumn environment. Changing its value should be done outside the multicols environment. Since `\tolerance` is set to `\multicoltolerance` at the beginning of every multicols environment one can locally overwrite this default by assigning `\tolerance␣=␣`⟨*desired value*⟩. There also exists a `\multicolpretolerance` parameter holding the value for `\pretolerance` within a multicols environment. Both parameters are usually used only by package designers.

Generation of multicolumn output can be divided into two parts. In the first part we are collecting material for a page, shipping it out, collecting material for the next page, and so on. As a second step, balancing will be done when the end of the multicols environment is reached. In the first step TeX might consider more material whilst finding the final columns than it actually use when shipping out the page. This might cause a problem if a footnote is encountered in the part of the input considered, but not used, on the current page. In

[2] This is dictated by lack of time. To implement floats one has to reimplement the whole LaTeX output routine.

[3] Actually the added space may be less because we use `\addvspace` (see the LaTeX manual for further information about this command).

[4] Look at the next paragraph, it was set with the `\sloppy` declaration.

this case the footnote might show up on the current page, while the footnotemark corresponding to this footnote might be set on the next one.[5] Therefore the multicols environment gives a warning message[6] whenever it is unable to use all the material considered so far.

If you don't use footnotes too often the chances of something actually going wrong are very slim, but if this happens you can help TeX by using a `\pagebreak` command in the final document. Another way to influence the behavior of TeX in this respect is given by the counter variable 'collectmore'. If you use the `\setcounter` declaration to set this counter to ⟨*number*⟩, TeX will consider ⟨*number*⟩ more (or less) lines before making its final decision. So a value of -1 may solve all your problems at the cost of slightly less optimal columns.

In the second step (balancing columns) we have other bells and whistles. First of all you can say `\raggedcolumns` if you don't want the bottom lines to be aligned. The default is `\flushcolumns`, so TeX will normally try to make both the top and bottom baselines of all columns align.

Additionally you can set another counter, the 'unbalance' counter, to some positive ⟨*number*⟩. This will make all but the right-most column ⟨*number*⟩ of lines longer than they would normally have been. 'Lines' in this context refer to normal text lines (i.e. one `\baselineskip` apart); thus, if your columns contain displays, for example, you may need a higher ⟨*number*⟩ to shift something from one column into another.

Unlike 'collectmore,' the 'unbalance' counter is reset to zero at the end of the environment so it only applies to one multicols environment.

The two methods may be combined but I suggest using these features only when fine tuning important publications.

Two more general tuning possibilities were added with version 1.5. TeX allows to measure the badness of a column in terms of an integer value, where 0 means optimal and any higher value means a certain amount of extra white space. 10000 is considered to be infinitely bad (TeX does not distinguish any further). In addition the special value 100000 means overfull (i.e., the column contains more text than could possibly fit into it).

The new release now measures every generated column and ignores solutions where at least one column has a badness being larger than the value of the counter columnbadness. The default value for this counter is 10000, thus TeX will accept all solutions except those being overfull. By setting the counter to a smaller value you can force the algorithm to search for solutions that do not have columns with a lot of white space.

However, if the setting is too low, the algorithm may not find any acceptable solution at all and will then finally choose the extreme solution of placing all text into the first column.

Often, when colunms are balanced, it is impossible to find a solution that distributes the text evenly over all columns. If that is the case the last column usually has less text than the others. In the earlier releases this text was stretched to produce a column with the same height as all others, sometimes resulting in really ugly looking columns.

In the new release this stretching is only done if the badness of the final column is not larger than the value of the counter finalcolumnbadness. The default setting is 9999, thus preventing the stretching for all columns that TeX would consider infinitely bad. In that case the final column is allowed to run short which gives a much better result.

2.2 Floats inside a multicols environment

Within the multicols environment the usual star float commands are available but their function is somewhat different as in the two-column mode of standard LaTeX. Stared floats, e.g., `figure*`, denote page wide floats that are handled in a similar fashion as normal floats outside the multicols environment. However, they will never show up on the page where they are encountered. In other words, one can influence their placement by specifying a combination of `t`, `b`, and/or `p` in their optional argument, but `h` doesn't work because the first possible place is the top of the next page. One should also note, that this means that their placement behavior is determined by the values of `\topfraction`, etc. rather then by `\dbl...`.

2.3 Warnings

Under certain circumstances the use of the multicols environment may result in in some warnings from TeX or LaTeX. Here is a

[5] The reason behind this behavior is the asynchronous character of the TeX *page_builder*. However, this could be avoided by defining very complicated output routines which don't use TeX primitives like `\insert` but do everything by hand. This is clearly beyond the scope of a weekend problem.

[6] This message will be generated even if there are no footnotes in this part of the text.

list of the important ones and the possible cause:

`Underfull \hbox (badness ...)`
: As the columns are often very narrow TeX wasn't able to find a good way to break the paragraph. Underfull denotes a loose line but as long the badness values is below 10000 the result is probably acceptable.

`Underfull \vbox ... while \output is active`
: If a column contains an character with an unusual depth, for example a '(', in the bottom line then this message may show up. It usually has no significance as long as the value is not more than a few points.

`LaTeX Warning: I moved some lines to the next page`
: As mentioned above, multicols sometimes screws up the footnote numbering. As a precaution, whenever there is a footnote on a page that where multicols had to leave a remainder for the following page this warning appears. Check the footnote numbering on this page. If it turns out that it is wrong you have to manually break the page using `\newpage` or `\pagebreak[..]`.

`Floats and marginpars not allowed inside 'multicols' environment!`
: This message appears if you try to use the `\marginpar` command or an unstared version of the figure or table environment. Such floats will disappear!

2.4 Tracing the output

To understand the reasoning behind the decisions TeX makes when processing a multicols environment, a tracing mechanism is provided. If you set the counter 'multicols' to a positive ⟨*number*⟩ you then will get some tracing information on the terminal and in the transcript file:

⟨*number*⟩ = 1. TeX will now tell you, whenever it enters or leaves a multicols environment, the number of columns it is working on and its decision about starting a new page before or after the environment.

⟨*number*⟩ = 2. In this case you also get information from the balancing routine: the heights tried for the left and right-most columns, information about shrinking if the `\raggedcolumns` declaration is in force and the value of the 'unbalance' counter if positive.

⟨*number*⟩ = 3. Setting ⟨*number*⟩ to this value will additionally trace the mark handling algorithm. It will show what marks are found, what marks are considered, etc. To fully understand this information you will probably have to read carefully trough the implementation.

⟨*number*⟩ ≥ 4. Setting ⟨*number*⟩ to such a high value will additionally place an `\hrule` into your output, separating the part of text which had already been considered on the previous page from the rest. Clearly this setting should *not* be used for the final output. It will also activate even more debugging code for mark handling.

3 Prefaces to older versions

3.1 Preface to version 1.4

Beside fixing some bugs as mentioned in the multicol.bug file this new release enhances the multicols environment by allowing for balancing in arbitrary contexts. It is now, for example, possible to balance text within a multicols or a minipage as shown in 2 where a multicols environment within a quote environment was used. It is now even possible to nest multicols environments.

The only restriction to such inner multicols environments (nested, or within TeX's internal vertical mode) is that such variants will produce a box with the balanced material in it, so that they can not be broken across pages or columns.

Additionally I rewrote the algorithm for balancing so that it will now produce slightly better results.

I updated the source documentation but like to apologize in advance for some 'left over' parts that slipped through the revision.

A note to people who like to improve the balancing algorithm of multicols: The balancing routine in now placed into a single macro which is called `\balance@columns`. This means that one can easily try different balancing routines by rewriting this macro. The interface for it is explained in table 1. There are several improvements possible, one can think of integrating the `\badness` function of TeX3, define a faster algorithm for find-

The macro \balance@columns that contains the code for balancing gathered material is a macro without parameters. It assumes that the material for balancing is stored in the box \mult@box which is a \vbox. It also “knows” about all parameters set up by the multicols environment, like \col@number, etc. It can also assume that \@colroom is the still available space on the current page.

When it finishes it must return the individual columns in boxes suitable for further processing with \page@sofar. This means that the left column should be stored in box register \mult@gfirstbox, the next in register \mult@firstbox $+ 2$, ..., only the last one as an exception in register \mult@grightbox. Furthermore it has to set up two the macros \kept@firstmark and \kept@botmark to hold the values for the first and bottom mark as found in the individual columns. There are some helper functions defined in section ?? which may be used for this. Getting the marks right “by hand” is non-trivial and it may pay off to first take a look at the documentation and implementation of \balance@columns below before trying anew.

Table 1: Interface description for \balance@columns

ing the right column height, etc. If somebody thinks he/she has an enhancement I would be pleased to learn about it. But please obey the copyright notice and don’t change multicol.dtx directly!

3.2 Preface to version 1.2

After the article about the multicols environment was published in *TUGboat* 10#3, I got numerous requests for these macros. However, I also got a changed version of my style file, together with a letter asking me if I would include the changes to get better paragraphing results in the case of narrow lines. The main differences to my original style option were additional parameters (like \multicoladjdemerits to be used for \adjdemerits, etc.) which would influence the line breaking algorithm.

But actually resetting such parameters to zero or even worse to a negative value won’t give better line breaks inside the multicols environment. TeXs line breaking algorithm will only look at those possible line breaks which can be reached without a badness higher than the current value of \tolerance (or \pretolerance in the first pass). If this isn’t possible, then, as a last resort, TeX will produce overfull boxes. All those (and only those) possible break points will be considered and finally the sequence which results in the fewest demerits will be chosen. This means that a value of -1000 for \adjdemerits instructs TeX to prefer visibly incompatible lines instead of producing better line breaks.

However, with TeX 3.0 it is possible to get decent line breaks even in small columns by setting \emergencystretch to an appropriate value. I implemented a version which is capable of running both in the old and the new TeX (actually it will simply ignore the new feature if it is not available). The calculation of \emergencystretch is probably incorrect. I made a few tests but of course one has have much more experience with the new possibilities to achieve the maximum quality.

Version 1.1a had a nice ‘feature’: the penalty for using the forbidden floats was their ultimate removal from LaTeXs \@freelist so that after a few \marginpars inside the multicols environment floats where disabled forever. (Thanks to Chris Rowley for pointing this out.) I removed this misbehaviour and at the same time decided to allow at least floats spanning all columns, e.g., generated by the figure* environment. You can see the new functionality in table 2 which was inserted at this very point. However single column floats are still forbidden and I don’t think I will have time to tackle this problem in the near future. As an advice for all who want to try: wait for TeX 3.0. It has a few features which will make life much easier in multicolumn surroundings. Nevertheless we are working here at the edge of TeXs capabilities, really perfect solutions would need a different approach than it was done in TeXs page builder.

The text below is nearly unchanged, I only added documentation at places where new code was added.

`\setemergencystretch`: This is a hook for people who like to play around. It is supposed to set the `\emergencystretch` ⟨*dimen*⟩ register provided in the new TeX 3.0. The first argument is the number of columns and the second one is the current `\hsize`. At the moment the default definition is `4pt` $\times$ `#1`, i.e. the `\hsize` isn't used at all. But maybe there are better formulae.

`\set@floatcmds`: This is the hook for the experts who like to implement a full float mechanism for the multicols environment. The `@` in the name should signal that this might not be easy.

Table 2: The new commands of multicol.sty version 1.2. Both commands might be removed if good solutions to these open problems are found. I hope that these commands will prevent that nearly identical style files derived from this one are floating around.