The \texttt{setouterhbox} package

Heiko Oberdiek\textsuperscript{*}

\texttt{<heiko.oberdiek at googlemail.com>}

2016/05/16 v1.8

Abstract

If math stuff is set in an \texttt{hbox}, then \TeX~performs some optimization and omits the implicit penalties \texttt{\binoppenalty} and \texttt{\relpenalty}. This package tries to put stuff into an \texttt{hbox} without getting lost of those penalties.

Contents

1 Documentation 2
   1.1 Introduction ................................. 2
   1.2 Acknowledgement .......................... 2
   1.3 Usage ........................................ 2
   1.4 Option \texttt{hyperref} .................... 3
   1.5 Example ..................................... 3

2 Implementation 4
   2.1 Package start stuff .......................... 4
   2.2 Interface macros ............................ 6
   2.3 Main part .................................... 6
   2.4 Environment support ........................ 8
   2.5 Option \texttt{hyperref} .................... 9

3 Test 9
   3.1 Catcode checks for loading ................ 9
   3.2 Test with package \texttt{url} ............... 11

4 Installation 11
   4.1 Download ..................................... 11
   4.2 Bundle installation .......................... 12
   4.3 Package installation .......................... 12
   4.4 Refresh file name databases ................ 12
   4.5 Some details for the interested ............ 12

5 Catalogue 13

6 References 13

7 History 14
   \[2005/10/05 v1.0\] .................................. 14
   \[2005/10/07 v1.1\] .................................. 14
   \[2005/10/18 v1.2\] .................................. 14
   \[2006/02/12 v1.3\] .................................. 14
   \[2006/08/26 v1.4\] .................................. 14
   \[2007/04/26 v1.5\] .................................. 14

\textsuperscript{*}Please report any issues at https://github.com/ho-tex/oberdiek/issues
1 Documentation

1.1 Introduction

There is a situation in hyperref’s driver for dvips where the user wants to have links that can be broken across lines. However dvips doesn’t support the feature. With option breaklinks hyperref sets the links as usual, put them in a box and write the link data with box dimensions into the appropriate \specials. Then, however, it does not set the complete unbreakable box, but it unwrappes the material inside to allow line breaks. Of course line breaking and glue setting will falsify the link dimensions, but line breaking was more important for the user.

1.2 Acknowledgement

Jonathan Fine, Donald Arsenaup and me discussed the problem in the newsgroup comp.text.tex where Damian Menscher has started the thread, see [1].

The discussion was productive and generated many ideas and code examples. In order to have a more permanent result I wrote this package and tried to implement most of the ideas, a kind of summary of the discussion. Thus I want and have to thank Jonathan Fine and Donald Arsenaup very much.

Two weeks later David Kastrup (posting in comp.text.tex, [2]) remembered an old article of Michael Downes ([3]) in TUGboat, where Michael Downes already presented the method we discuss here. Nowadays we have $\varepsilon$-TEX that extends the tool set of a T\TeX macro programmer. Especially useful $\varepsilon$-TEX was in this package for detecting and dealing with erroneous situations.

However also nowadays a perfect solution for the problem is still missing at macro level. Probably someone has to go deep in the internals of the T\TeX compiler to implement a switch that let penalties stay where otherwise T\TeX would remove them for optimization reasons.

1.3 Usage

Package loading. $\LaTeX$: as usually:

\usepackage{setouterhbox}

The package can also be included directly, thus plain T\TeX users write:

\input setouterhbox.sty

Register allocation. The material will be put into a box, thus we need to know these box number. If you need to allocate a new box register:

$\LaTeX$: \newsavebox{\langle name\rangle}

plain T\TeX: \newbox{\langle name\rangle}

Then \langle name\rangle is a command that held the box number.
Box wrapping. \LaTeX users put the material in the box with an environment similar to \texttt{lrbox}. The environment \texttt{setouterhbox} uses the same syntax and offers the same features, such as verbatim stuff inside:

\begin{setouterhbox}\{⟨box number⟩\}\end{setouterhbox}

Users with plain \TeX do not have environments, they use instead:

\setouterhbox{⟨box number⟩}\endsetouterhbox

In both cases the material is put into an \texttt{hbox} and assigned to the given box, denoted by \texttt{⟨box number⟩}. Note the assignment is local, the same way \texttt{lrbox} behaves.

Unwrapping. The box material is ready for unwrapping:

\unhbox ⟨box number⟩

1.4 Option hyperref

Package \texttt{url} uses math mode for typesetting urls. Break points are inserted by \texttt{\binoppenalty} and \texttt{\relpenalty}. Unhappily these break points are removed, if \texttt{hyperref} is used with option breaklinks and drivers that depend on \texttt{pdfmark}: \texttt{dvips}, \texttt{vtexpdfmark}, \texttt{textures}, and \texttt{dvipsone}. Thus the option \texttt{hyperref} enables the method of this package to avoid the removal of \texttt{\relpenalty} and \texttt{\binoppenalty}. Thus you get more break points. However, the link areas are still wrong for these drivers, because they are not supporting broken links.

Note, you need version 2006/08/16 v6.75c of package \texttt{hyperref}, because starting with this version the necessary hook is provided that package \texttt{setouterhbox} uses.

\usepackage[...]{hyperref}[2006/08/16]
\usepackage[hyperref]{setouterhbox}

Package order does not matter.

1.5 Example

1 \{∗example\}
2 \documentclass[a5paper]{article}
3 \usepackage[url][2005/06/27]
4 \usepackage[setouterhbox]
5
6 \newsavebox{\testbox}
7
8 \setlength{\parindent}{0pt}
9 \setlength{\parskip}{2em}
10
11 \begin{document}
12 \raggedright
13
14 \url{http://this.is.a.very.long.host.name/followed/%
15 by/a/very_long_long_long_path.html}\
16
17 \abox{\testbox}\%
18 \url{http://this.is.a.very.long.host.name/followed/%
19 by/a/very_long_long_long_path.html}\
20 \%}
21 \unhbox{\testbox}
22
23 \begin{setouterhbox}{\testbox}\%
24 \url{http://this.is.a.very.long.host.name/followed/%
25 by/a/very_long_long_long_path.html}%
2 Implementation

Internal macros are prefixed by \setouterhbox, \@ is not used inside names, thus we do not need to care of its catcode if we are not using it as \LaTeX{} package.

2.1 Package start stuff

Prevent reloading more than one, necessary for plain \TeX{}: Reload check, especially if the package is not used with \LaTeX{}.

\begin{verbatim}
\begingroup\catcode61\catcode48\catcode32=10\relax%
\catcode13=5 \^M
\endlinechar=13 \%
\catcode35=6 \#'
\catcode39=12 \'
\catcode44=12 ,
\catcode45=12 -
\catcode46=12 .
\catcode48=12 :
\catcode64=11 \@
\catcode123=1 \{
\catcode125=2 \}
\expandafter\let\expandafter\x\csname ver@setouterhbox.sty\endcsname
\ifx\x\relax % plain-TeX, first loading
\else
\def\empty{}%
\ifx\x\empty % LaTeX, first loading, % variable is initialized, but \ProvidesPackage not yet seen
\else
\expandafter\ifx\csname PackageInfo\endcsname\relax
\x#1#2{%
\immediate\write-1{Package #1 Info: #2.}%
}
\else
\def\x#1#2{\PackageInfo{#1}{#2, stopped}}%
\fi
\fi
\aftergroup\endinput
\fi
\endgroup%
\end{verbatim}

Package identification:

\begin{verbatim}
\begingroup\catcode61\catcode48\catcode32=10\relax%
\catcode13=5 \^M
\endlinechar=13 \%
\catcode35=6 \#
\catcode39=12 \'
\catcode40=12 (
\catcode41=12 )
\catcode44=12 ,
\catcode45=12 -
\catcode46=12 .
\catcode47=12 /
\catcode48=12 :
\catcode64=11 \@
\end{verbatim}
\setouterhboxBox The method requires a global box assignment. To be on the safe side, a new box register is allocated for this global box assignment.

\setouterhboxFailure Error message for both plain \TeX and \LaTeX.

\setouterhboxRemove Remove all kern, glue, and penalty nodes; poor man’s version, if \e\TeX is not available.

\setouterhbox Passing the box contents by macro parameter would prevent catcode changes in the box contents like by \verb!\bgroup! and \verb!\egroup! does not work, because stuff has to be added at the begin and end of the box, thus the syntax \setouterhbox{⟨box number⟩}…\endsetouterhbox is used. Also we automatically get an environment \setouterhbox if \e\TeX is used.
Most of the work is done in the end part, thus the heart of the method follows:

\def\endsetouterhbox{% \endgroup 
Omit the first pass to get the penalties of the second pass.
\pretolerance-1 %
\tolerance10000 %
\hsize\maxdimen

Line is not underfull:
\parfillskip 0pt plus 1fill\relax
\leftskip0pt\relax
Suppress underful \hbox warnings, is explicit line breaks are used.
\rightskip0pt plus 1fill\relax
\everypar{%
Ensure that there is a paragraph and prevents \endgraf from eating terminal glue:
\kern0pt%
\endgraf
\setouterhboxRemove
\ifnum\lastnodetype=1 %
\loop
\setouterhboxRemove
\ifnum\lastnodetype=1 %
\setbox0=\lastbox
\global\setbox\setouterhboxBox=\hbox{\unhbox0 % Remove \rightskip, a penalty with -10000 is part of the previous line.}
\else
\setouterhboxFailure{%
\MessageBreak
(\string\lastnodetype: \number\lastnodetype, expected: 1)%
}\fi
\setouterhboxRemove
\loop
\setouterhboxRemove
\ifcase\ifvoid0 1\else0\fi
\global\setbox\setouterhboxBox=\hbox{%
\unhbox0 %
Remove \rightskip, a penalty with -10000 is part of the previous line.
There was just one line that we have caught.  
\[\text{Expected stuff was detected before the line.}\]

2.4 Environment support

Check \texttt{\@currenvir} for the case that \texttt{\setouterhbox} was called as environment. Then the box assignment must be put after the \texttt{\endgroup} of \texttt{\end{…}}.

\begin{verbatim}
258 \def\setouterhboxLast#1{\setbox#1\hbox{\unhbox\setouterhboxBox \unskip % remove \rightskip glue \unskip % remove \parfillskip glue \unpenalty % remove \penalty 10000 \unkern % remove explicit kern inserted above} \setouterhboxFinish#1 is an explicit number.  
\end{verbatim}
\setouterhboxAfter \#1 is an explicit number.
278 \def\setouterhboxAfter#1#2\NIL{% 279 \aftergroup#1\%
280 \ifx\#2\%
281 \else
282 \setouterhboxReturnAfterFi{% 283 \setouterhboxAfter#2\NIL 284 }\%
285 \fi
286 }

\setouterhboxReturnAfterFi A utility macro to get tail recursion.
287 \long\def\setouterhboxReturnAfterFi#1\fi{\fi#1} 288 \setouterhboxReturnAfterFi

Restore catcodes we have need to distinguish between the implementation with and without \$\varepsilon$-\TeX.
289 \catcode69=11\relax % E
290 \catcode84=11\relax % T

2.5 Option hyperref
290 \begingroup
291 \def\x{LaTeX2e}\%
292 \expandafter\endgroup
293 \ifx\x\fmtname
294 \else
295 \expandafter\setouterhboxAtEnd
296 \fi%
297 \Hy@setouterhbox \Hy@setouterhbox is the internal hook that hyperref uses since 2006/02/12 v6.75a.
298 \DeclareOption{hyperref}{% 299 \long\def\Hy@setouterhbox#1#2{% 300 \setouterhbox{#1}#2\endsetouterhbox 301 }% 302 }
303 \ProcessOptions\relax
304 \setouterhboxAtEnd%
305 (/package)

3 Test
3.1 Catcode checks for loading
305 (*test1)
306 \catcode`\{=1 %
307 \catcode`\} =2 % 308 \catcode`\#=6 %
309 \catcode`@=11 %
310 \expandafter\ifx\csname count@\endcsname\relax
311 \countdef\count@=255 %
312 \fi
313 \expandafter\ifx\csname gobble\endcsname\relax
314 \long\def\gobble#1{% 315 \fi
316 \expandafter\ifx\csname @firstofone\endcsname\relax
317 \long\def\@firstofone#1{% 318 \fi
319 \expandafter\ifx\csname loop\endcsname\relax
320 \expandafter\@firstofone
321 \else
\def\loop#1\repeat{\def\body{#1}\iterate}
\def\iterate{\body\let\next\iterate\else\let\next\relax\fi\next}
\let\repeat=\fi
\def\RestoreCatcodes{}
\count@=0 \loop\edef\RestoreCatcodes{\RestoreCatcodes\catcode\the\count@=\the\catcode\count@\relax}\ifnum\count@<255 \advance\count@ 1 \repeat}
\def\RangeCatcodeInvalid#1#2{\count@=#1\relax\loop\catcode\count@=15 \ifnum\count@<#2\relax\advance\count@ 1 \repeat}
\def\RangeCatcodeCheck#1#2#3{\count@=#1\relax\loop\ifnum#3=\catcode\count@\else\errmessage{Character \the\count@ with wrong catcode \the\catcode\count@ instead of \number#3} \fi\ifnum\count@<#2\relax\advance\count@ 1 \repeat}
\def\space{ }\expandafter\ifx\csname LoadCommand\endcsname\relax\def\LoadCommand{\input setouterhbox.sty}\relax\fi
\def\Test{\RangeCatcodeInvalid{0}{47}\RangeCatcodeInvalid{58}{64}\RangeCatcodeInvalid{91}{96}\RangeCatcodeInvalid{123}{255}\catcode`\@=12 \catcode`\=0 \space( }
\LoadCommand
\RangeCatcodeCheck{0}{36}{15} %
\RangeCatcodeCheck{37}{37}{14} %
\RangeCatcodeCheck{38}{47}{15} %
\RangeCatcodeCheck{48}{57}{12} %
\RangeCatcodeCheck{58}{63}{15} %
\RangeCatcodeCheck{64}{64}{12} %
\RangeCatcodeCheck{65}{90}{11} %
\RangeCatcodeCheck{91}{91}{15} %
\RangeCatcodeCheck{92}{92}{0} %
\RangeCatcodeCheck{93}{96}{15} %
\RangeCatcodeCheck{97}{122}{11} %
\RangeCatcodeCheck{123}{255}{15} %
\RestoreCatcodes
}
\Test
\csname @@end\endcsname
\end
⟨/test1⟩

3.2 Test with package url

⟨*/test2⟩
\nofiles
\documentclass[a5paper]{article}
\usepackage[url]{url}[2005/06/27]
\usepackage{setouterhbox}
\newsavebox\testbox
\setlength{\parindent}{0pt}
\setlength{\parskip}{2em}
\begin{document}
\raggedright
\url{http://this.is.a.very.long.host.name/followed/%
by/a/very_long_long_path.html} %
\sbox\testbox{\url{http://this.is.a.very.long.host.name/followed/%
by/a/very_long_long_path.html} %}
\unhbox\testbox
\begin{setouterhbox}{\testbox} %
\url{http://this.is.a.very.long.host.name/followed/%
by/a/very_long_long_path.html} %
\end{setouterhbox}
\unhbox\testbox
\end{document}
⟨/test2⟩

4 Installation

4.1 Download

Package. This package is available on CTAN\(^1\):

\(^1\)http://ctan.org/pkg/setouterhbox
**Bundle.** All the packages of the bundle ‘oberdiek’ are also available in a TDS compliant ZIP archive. There the packages are already unpacked and the documentation files are generated. The files and directories obey the TDS standard.

*TDS* refers to the standard “A Directory Structure for T\TeX\ Files” (CTAN:tds/tds.pdf). Directories with `texmf` in their name are usually organized this way.

## 4.2 Bundle installation

**Unpacking.** Unpack the `oberdiek.tds.zip` in the TDS tree (also known as `texmf` tree) of your choice. Example (linux):

```
unzip oberdiek.tds.zip -d ~/texmf
```

**Script installation.** Check the directory `TDS:scripts/oberdiek/` for scripts that need further installation steps. Package `attachfile2` comes with the Perl script `pdfatfi.pl` that should be installed in such a way that it can be called as `pdfatfi`. Example (linux):

```
chmod +x scripts/oberdiek/pdfatfi.pl
cp scripts/oberdiek/pdfatfi.pl /usr/local/bin/
```

## 4.3 Package installation

**Unpacking.** The `.dtx` file is a self-extracting `docstrip` archive. The files are extracted by running the `.dtx` through plain `T\TeX\`:

```
tex setouterhbox.dtx
```

**TDS.** Now the different files must be moved into the different directories in your installation TDS tree (also known as `texmf` tree):

```
setouterhbox.sty → tex/generic/oberdiek/setouterhbox.sty
setouterhbox.pdf → doc/latex/oberdiek/setouterhbox.pdf
setouterhbox-example.tex → doc/latex/oberdiek/setouterhbox-example.tex
test/setouterhbox-test1.tex → doc/latex/oberdiek/test/setouterhbox-test1.tex
test/setouterhbox-test2.tex → doc/latex/oberdiek/test/setouterhbox-test2.tex
setouterhbox.dtx → source/latex/oberdiek/setouterhbox.dtx
```

If you have a `docstrip.cfg` that configures and enables `docstrip`'s TDS installing feature, then some files can already be in the right place, see the documentation of `docstrip`.

## 4.4 Refresh file name databases

If your `T\TeX\` distribution (te\TeX, mikt\TeX, ...) relies on file name databases, you must refresh these. For example, te\TeX users run `texhash` or `mktexlsr`.

## 4.5 Some details for the interested

**Unpacking with `\LaTeX`.** The `.dtx` chooses its action depending on the format:

- plain `T\TeX`: Run `docstrip` and extract the files.
- `\LaTeX`: Generate the documentation.

If you insist on using `\LaTeX` for `docstrip` (really, `docstrip` does not need `\LaTeX`), then inform the autodetect routine about your intention:

```
latex \let\install=y\input{setouterhbox.dtx}
```

Do not forget to quote the argument according to the demands of your shell.
Generating the documentation. You can use both the .dtx or the .drv to generate the documentation. The process can be configured by the configuration file ltxdoc.cfg. For instance, put this line into this file, if you want to have A4 as paper format:

\PassOptionsToClass{a4paper}{article}

An example follows how to generate the documentation with pdflatex:\n
\begin{verbatim}
<latex syntax=``>
\PassOptionsToClass{a4paper}{article}
\sloppy
\begin{verbatim}
pdflatex setouterhbox.dtx
makeindex -s gind.ist setouterhbox.idx
pdflatex setouterhbox.dtx
makeindex -s gind.ist setouterhbox.idx
pdflatex setouterhbox.dtx
\end{verbatim}
\end{verbatim}

5 Catalogue

The following XML file can be used as source for the TeX Catalogue. The elements caption and description are imported from the original XML file from the Catalogue. The name of the XML file in the Catalogue is setouterhbox.xml.

\begin{verbatim}
<?xml version='1.0' encoding='us-ascii'?>
<!DOCTYPE entry SYSTEM 'catalogue.dtd'>
<entry datestamp='$Date$' modifier='$Author$' id='setouterhbox'>
  <name>setouterhbox</name>
  <caption>Set hbox in outer horizontal mode.</caption>
  <authorref id='auth:oberdiek'/>
  <copyright owner='Heiko Oberdiek' year='2005-2007'/>
  <license type='lppl1.3'/>
  <version number='1.8'/>
  <description>
    If math stuff is set in an \hbox\<tt>, then \TeX\ performs some optimization and omits the implicit penalties \binoppenalty\ and \relpenalty\.<tt>. This package tries to put stuff into an \hbox\<tt> without losing those penalties.
  </description>
  <documentation details='Package documentation' href='ctan:/macros/latex/contrib/oberdiek/setouterhbox.pdf'/>
  <ctan file='true' path='/macros/latex/contrib/oberdiek/setouterhbox.dtx'/>
  <miktex location='oberdiek'/>
  <texlive location='oberdiek'/>
  <install path='/macros/latex/contrib/oberdiek/oberdiek.tds.zip'/>
</entry>
\end{verbatim}

6 References


7 History

[2005/10/05 v1.0]
• First version.

[2005/10/07 v1.1]
• Option hyperref added.

[2005/10/18 v1.2]
• Support for explicit line breaks added.

[2006/02/12 v1.3]
• DTX format.
• Documentation extended.

[2006/08/26 v1.4]
• Date of hyperref updated.

[2007/04/26 v1.5]
• Use of package infwarerr.

[2007/05/17 v1.6]
• Standard header part for generic files.

[2007/09/09 v1.7]
• Catcode section added.

[2016/05/16 v1.8]
• Documentation updates.

8 Index

Numbers written in italic refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; plain numbers refer to the code lines where the entry is used.

<table>
<thead>
<tr>
<th>Symbols</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>308</td>
<td>\advance</td>
</tr>
<tr>
<td>%</td>
<td>384</td>
<td>\aftergroup</td>
</tr>
<tr>
<td>@</td>
<td>309, 382</td>
<td>@firstofone</td>
</tr>
<tr>
<td>@gobble</td>
<td>314, 322</td>
<td>@undefined</td>
</tr>
<tr>
<td>\</td>
<td>280, 383</td>
<td>\begin</td>
</tr>
<tr>
<td>{</td>
<td>306</td>
<td>\binoppenalty</td>
</tr>
<tr>
<td>}</td>
<td>307</td>
<td>\body</td>
</tr>
</tbody>
</table>