The collref Package*

Niklas Beisert

Max-Planck-Institut für Gravitationsphysik
Albert-Einstein-Institut Am Mühlenberg 1
14476 Potsdam, Germany

nbeisert@aei.mpg.de

9 June 2009, v1.0

Abstract

collref is a LaTeX 2_ε package to automatically collect multiple `\bibitem` references which always appear in the same sequence in `\cite` into a single `\bibitem` block.

Contents

1 Introduction

Suppose a manuscript uses the following set of four references:

[1] Reference A

[2] Reference B

[3] Reference C

[4] Reference D

Now if references B and C cover similar or related material, they might always be cited together as in "[..., 2, 3, ...]" throughout the manuscript. In some (physics) journals it is then customary to collect the two references into a single reference

[1] Reference A

[2] Reference B
Reference C

*AEI-2009-054

[3] Reference D

and cite it by "[..., 2, ...]". The package collref automates this process by analysing the `\cite` commands and identifying blocks of references which always appear in conjunction. These blocks are collapsed to a single item in the bibliography. Please note that collref requires the sequence of `\bibitem` entries to match with the sequence of `\cite` blocks. This is most easily achieved through the use of BibTeX with any *unsorted* style.

Similar CTAN Packages. The objective and some of the implementation of the collref package is similar to the CTAN packages mcite by Thorsten Ohl and mciteplus by Michael Shell, but the functionality is different is several respects:

- collref is intended to work transparently: LaTeX documents which compile with collref should also compile fine without invoking collref (obviously without collected references). The package decides automatically which references can be collapsed, no further interaction of the author is required.

  mcite and mciteplus leave the decision/duty to collapse certain references using the modified syntax `\cite{A,*B,*C}`.

- mcite and mciteplus are intended to handle punctuations in collapsed references correctly. This requires a specialised BibTeX style.

  No effort is made in collref in this regard. Some minor modification in `collref.sty` together with a modified BibTeX style might achieve basic punctuation features similar to mcite.

2 Usage

Inclusion. To use collref simply add the command

```
\usepackage{collref}
```

to the preamble of your LaTeX document. No further interaction is required.

Bibliography Preparation. Please note that only such blocks of references can be collapsed which appear in the same order for `\cite` commands as for `thebibliography`. It is recommended to prepare the bibliography through BibTeX which does this automatically. You must use a style which does not sort the references but preserves the order in which they were `\cite`'d, e.g. `unsrt.bst`.

Separators. collref does not provide correct linguistic punctuation between collected references, but it allows to specify how collected references are separated. This is specified through the package options `\usepackage[`*opt*`]{collref}` where *opt* is one of the following

`nosep` (default)

no separator:

[1] A
[2] B C
[3] D

`parsep`

separated by `\par`:

[1] A

[2] B

C

[3] D

`bulletsep`

separated by •:

[1] A
[2] B • C
[3] D

Alternative separators can be specified in the preamble through the command:

```
\collectsep{separator}
```

Control. The package collref provides one command to control which references (not) to collect:

\nocollect{*label*}

It ensures that the label *label* starts a new \bibitem. It is not collapsed with earlier references. Later references, however, can still be collapsed to the end of *label*.

Interaction with CTAN Packages. The package collref has been tested with other CTAN packages concerned with the citations and the bibliography:

- cite: collref works in conjunction with cite. Note that you must load cite *before* collref so that the latter can pass the correctly reduced list of references down to cite. Tested with cite v5.1.
- hyperref: collref works in conjunction with hyperref. The two packages can be loaded in any sequence. Tested with hyperref v6.78s.

A Files and Installation

The package consists of the files

README	readme file
collref.ins	installation file
collref.dtx	source file
collref.sty	package file
collsamp.tex	sample file
collref.pdf	manual

The distribution consists of the files README, collref.ins and collref.dtx.

- Run (pdf)LaTeX on collref.dtx to compile the manual collref.pdf (this file).
- Run LaTeX on collref.ins to create the package collref.sty and the sample collsamp.tex. Copy the file collref.sty to an appropriate directory of your LaTeX distribution, e.g. *texmf-root*/tex/latex/collref.

B Sample File

In this section we provide a sample file.

```
\documentclass{article}
%\usepackage{cite}
\usepackage[bulletsep]{collref}
%\usepackage{hyperref}

\begin{document}

\cite{c1,c2,c3,c4}
\nocollect{c3}
\cite{c5,c6,c7,c8,c9}
\cite{c5,c6,c7}
\cite{c7,c8,c9}

\begin{thebibliography}{11}
\bibitem{c1} reference 1
\bibitem{c2} reference 2
```

```
\bibitem{c3} reference 3
\bibitem{c4} reference 4
\bibitem{c5} reference 5
\bibitem{c6} reference 6
\bibitem{c7} reference 7
\bibitem{c8} reference 8
\bibitem{c9} reference 9
\end{thebibliography}

\end{document}
```

It produces the output:

[1, 2] [3, 4, 5] [3, 4] [4, 5]

[1] reference 1 • reference 2

[2] reference 3 • reference 4

[3] reference 5 • reference 6

[4] reference 7

[5] reference 8 • reference 9

C Implementation

In this section we describe the package `collref.sty`.

Internal Lists. For each bibliography label *label* the package maintains a predecessor `\nc@p@`*label* and a successor `\nc@s@`*label*. These are initially undefined. When a label *label* is first cited these labels are set to the *predecessor* and *successor* labels, respectively, in `\cite{`*..., predecessor, label, successor, ...*`}`. An empty `\nc@p@`*label* or `\nc@s@`*label* refers to the beginning and end of a block, respectively. Whenever `\cite` finds conflicting blocks (non-matching predecessors or successors in two `\cite`'s), it terminates the blocks to the maximum common overlap.

Interface. The package provides two public commands, described above:

```
\newcommand{\collectsep}[1]{\def\nc@sep{#1}}
\newcommand{\nocollect}[1]{\nc@breakbefore{#1}}
```

Package Options. The package provides three predefined separators, described above:

```
\DeclareOption{nosep}{\collectsep{}}
\DeclareOption{parsep}{\collectsep{\par}}
\DeclareOption{bulletsep}{\collectsep{$\bullet$ }}
\ExecuteOptions{nosep}
\ProcessOptions
```

Internal Commands. Some internal commands for abbreviation:

```
\newcommand{\nc@getcsname}[1]{\csname #1\endcsname}
\newcommand{\nc@setcsname}[2]{\expandafter\xdef\csname #1\endcsname{#2}}
```

Command to terminate the chain before a label: The predecessor of the label is terminated. If the predecessor was active, its successor is also terminated.

```
\newcommand{\nc@breakbefore}[1]{%
  \edef\nc@citepred{\@ifundefined{nc@p@#1}{}{\nc@getcsname{nc@p@#1}}}%
  \ifx\nc@citepred\@empty\else\nc@setcsname{nc@s@\nc@citepred}{}\fi%
  \nc@setcsname{nc@p@#1}{}%
}
```

Command to terminate the chain after a label. Similar to the above command.

```
\newcommand{\nc@breakafter}[1]{%
  \edef\nc@citesucc{\@ifundefined{nc@s@#1}{}{\nc@getcsname{nc@s@#1}}}%
  \ifx\nc@citesucc\@empty\else\nc@setcsname{nc@p@\nc@citesucc}{}\fi%
  \nc@setcsname{nc@s@#1}{}%
}
```

Citations. Hack for `\@citex`: It is assumed that (as in LaTeX 2_ε) `\cite` eventually passes down to `\@citex`.

```
\let\nc@old@citex\@citex

\def\@citex[#1]#2{%
  \let\nc@citecomma\@empty%
  \let\nc@citestring\@empty%
  \let\nc@citelast\@empty%
```

Main loop to process the arguments of `\cite`. The current label is stored in `\nc@citethis`.

```
  \@for\nc@citethis:=#2\do{%
    \edef\nc@citethis{\expandafter\@firstofone\nc@citethis\@empty}%
```

The first entry has no predecessor, terminate the chain.

```
    \ifx\nc@citelast\@empty%
      \nc@breakbefore{\nc@citethis}%
    \else%
```

Non-first entry. Fill undefined successor and predecessors entries with the current chain sequence.

```
      \@ifundefined{nc@s@\nc@citelast}%
                   {\nc@setcsname{nc@s@\nc@citelast}{\nc@citethis}}{}%
      \@ifundefined{nc@p@\nc@citethis}%
                   {\nc@setcsname{nc@p@\nc@citethis}{\nc@citelast}}{}%
```

Get the successor and predecessors for the last and current entry, respectively.

```
      \edef\nc@citesucc{\nc@getcsname{nc@s@\nc@citelast}}%
      \edef\nc@citepred{\nc@getcsname{nc@p@\nc@citethis}}%
```

In case of mismatching chains: terminate all links.

```
      \ifx\nc@citesucc\nc@citethis%
        \ifx\nc@citepred\nc@citelast%
        \else%
          \nc@breakafter{\nc@citelast}%
          \nc@breakbefore{\nc@citethis}%
        \fi%
      \else%
        \nc@breakafter{\nc@citelast}%
        \nc@breakbefore{\nc@citethis}%
      \fi%
    \fi%
```

Get content of \b@*label* entry to find out whether the \bibitem{label} entry exists. We need to take special care of extended label definitions in hyperref.

```
74      {\def\hyper@@link[##1]##2##3##4{##4}%
75        \xdef\nc@citelabel{\nc@getcsname{b@\nc@citethis}}}%
```

Only add those labels which actually exist to the pass-on string. This removes collaped references from the citation marks.

```
76      \ifx\nc@citelabel\@empty\else%
77        \edef\nc@citestring{\nc@citestring\nc@citecomma\nc@citethis}%
78      \fi%
```

Write \citation tag to .aux file in original order. Some duplicate \citation's will be written by the original \citex code, but these will have no impact.

```
79      \if@filesw\immediate\write\@auxout{\string\citation{\nc@citethis}}\fi%
```

Continue to next label

```
80      \edef\nc@citelast{\nc@citethis}%
81      \def\nc@citecomma{,}%
82    }%
```

The last entry has no successor, terminate the chain.

```
83    \nc@breakafter{\nc@citelast}%
```

Pass on to original LaTeX code.

```
84    \nc@old@citex[#1]{\nc@citestring}%
85  }
```

Bibliography. Enhance the thebibliography environment to a) reset the \nc@biblast label to something, and b) convert linebreaks into whitespaces (avoid implicit \par's)

```
86 \let\nc@old@thebibliography\thebibliography
87
88 \def\thebibliography{%
89   \xdef\nc@biblast{asldjfhasklfh}%
90   \catcode`\^^M=10%
91 \nc@old@thebibliography}
```

Overwrite \bibitem: It is assumed that the native LaTeX 2_ε code is equivalent but with the LaTeX internals \@lbibitem and \@bibitem. Some other packages may also redefine \bibitem and this will inevitable cause compatibility issues. This implementation is safe with current versions of hyperref.

```
92 \def\bibitem{\@ifnextchar[\nc@lbibitem\nc@bibitem}
```

\nc@noitem is invoked in place of the original \@bibitem or \@lbibitem for collapsed references:

```
93 \def\nc@noitem#1{%
94   \nc@sep%
95   \if@filesw\immediate\write\@auxout{\string\bibcite{#1}{}}\fi%
96 \ignorespaces}
```

The hack for @bibitem: It checks whether this reference is part of a block. If so, collect by \nc@noitem, otherwise pass down to \@bibitem

```
97 \def\nc@bibitem#1{%
```

```
  \edef\nc@bibpred{\@ifundefined{nc@p@#1}{}{\nc@getcsname{nc@p@#1}}}%
  \ifx\nc@biblast\nc@bibpred\nc@noitem{#1}\else\@bibitem{#1}\fi%
  \xdef\nc@biblast{#1}%
\ignorespaces}
```

Similar hack for `@lbibitem`:

```
\def\nc@lbibitem[#1]#2{%
  \edef\nc@bibpred{\@ifundefined{nc@p@#2}{}{\nc@getcsname{nc@p@#2}}}%
  \ifx\nc@biblast\nc@bibpred\nc@noitem{#2}\else\@lbibitem[#1]{#2}\fi%
  \xdef\nc@biblast{#2}%
\ignorespaces}
```

D Copyright