1 Introduction

The \texttt{rkeyval} package provides functions for scanning key-value notation similar to the kind of scanning supported by the standard \texttt{keyval} package. However, the syntax is more restrictive in order to make some improved error-checking possible. In particular, if a comma is omitted between two instances of \texttt{key}={\textit{value}} form, the \texttt{RestrictedSetKeys} command will spot the missing comma and issue a suitable error message (and it will be given at the point where the missing comma is detected, before reading any further in the \TeX{} file). The standard \texttt{setkeys} command, by contrast, will append the second key name to the value of the first key and discard the second value, without any notification to the user that anything has gone wrong. But that is partly because the standard \texttt{setkeys} command allows implied values and does not require braces around explicit values (except when necessary to hide material that has a syntactic resemblance to a key-value pair). With \texttt{RestrictedSetKeys} the value must always be present and it must be enclosed in braces.

Further restrictions of the \texttt{RestrictedSetKeys} command and its companion commands reduce memory consumption in certain ways. Defining a key creates only one control sequence, a container for holding the value. Processing of key values is normally limited to storing a value given by the user; any additional processing must be supplied separately by the programmer.

Generally speaking, the error-checking done by \texttt{RestrictedSetKeys} is better for applications where all the keys are expected to have textual values, while
\setkeys is better when one wants to silently recover as far as possible from syntactic errors, instead of notifying the user of the errors; or when keys have nontrivial default values (i.e., not empty) or other kinds of special processing.

\RestrictedSetKeys{setup-code}{group}{code}{key={val}, ...}

Normally \RestrictedSetKeys simply carries out the following assignment for each key-value pair:
\def\group\'key{val}

The first argument is normally empty, but the exact nature of the warnings given and other aspects of the processing can be affected by putting extra setup code there. The amsrefs package uses this to implement a copying operation where field name and value are written out immediately to another file instead of being stored in the usual way.

Some examples for defining the key names associated with a given group.
This defines “title” as a recognized key for the bib group:
\DefineSimpleKey{bib}{title}

If a key is defined with \DefineSimpleKey, the result of using the same key more than once in a single entry will be an error message.

This defines “title” to be a repeatable key:
\DefineSupersedingKey{bib}{title}

If it occurs more than once, the last value supersedes the earlier ones, instead of getting an error. This variation is not needed for simple usage, but in more complicated situations where key values are combined from multiple sources, it may be useful.

This defines “author” to be a repeatable key, with each value being appended to a list of values:
\DefineAdditiveKey{bib}{author}{\name}

The third argument specifies a wrapper function that should be applied to each item in the list. I.e., suppose that two author names are given:

\begin{itemize}
    \item author={Smith, J. Q.},
    \item author={Taylor, X. Y.},
\end{itemize}

Then they will be stored in the form
\begin{itemize}
    \item name=Smith, J. Q.,
    \item name=Taylor, X. Y.,
\end{itemize}

This defines “transition” to be a dummy key with a value that is superficially nonempty but effectly empty:
\DefineDummyKey{bib}{transition}

Defining a dummy key like this can be useful in dealing with certain boundary situations that sometimes arise.
2 Implementation

Standard declaration of package name and date.
\NeedsTeXFormat{LaTeX2e}
\ProvidesPackage{rkeyval}[2004/05/05 v1.08]
\@xp
\@nx
\let\@xp\expandafter
\let\@nx\noexpand
\@emptytoks
Using \@ifundefined here avoids problems with really old versions of \LaTeX{} that choke on \texttt{newtoks} if it is written directly in the false branch of a conditional.
\@ifundefined{@emptytoks}{\csname newtoks\endcsname\@emptytoks}{}
\@emptokenb
\@ifundefined{@emptokenb}{\csname newtoks\endcsname\@emptokenb}{}
\@append
\def\@append#1#2#3{\@xp\def\@xp#2\@xp{#2#1{#3}}}
\star@
Test for a trailing option marked by a star. Usage:
\newcommand{\blub}[1]{\star@{\blubaux{#1}}{default}}
Arg 1 of \texttt{\star@} is the code to be run, arg 2 is the default value of the option (could be empty). If arg 1 is \texttt{\moo}, this test discards a star and expands to \texttt{\moo} if a star is found, or expands to \texttt{\moo{#2}} if not. As the example shows, arg 1 need not be a single token.
\def\star@#1#2{
\def\star@a##1{#1}
\def\star@b{#1{#2}}
\futurelet\@let@token\star@test}
\star@test{\ifx*\@let@token \let\star@b\star@a\fi\star@b}
Please note: If there is a space before the star, then the star is not treated as an option char.
Why use a star? Since it’s already part of standard \LaTeX{} command syntax, it’s unlikely to suffer from catcode changes.
Why not just put the star at the beginning in the usual way? It seemed to me that the lack of a trailing option feature was a deficiency in current \LaTeX{} and could be given an experimental implementation in a package like this without any adverse effect on existing documents.
Ensure non-weird catcode for relevant characters.
\edef\NormalCatcodes{elax}
\PushCatcodes\NormalCatcodes
\extract@group  Extracts “group” from \group'field.
\def\extract@group@a#1#2'{#2\nilgobble}

\section{Data structures}

The result of scanning the key/value pairs is an assignment statement for \rsk@toks. For example, consider the entry
\bib{LelekZ1962}{article}{
  author={Lelek, A.},
  author={Zaremba, D.},
  title={Dimensions of irreducible ...},
  journal={Fund. Math.},
  date={1962/63},
}

The scanned result is to assign
\global\rsk@toks{%
  \set:bib'author{Lelek, A.}{}
  \set:bib'author{Zaremba, D.}{}
  \set:bib'title{Dimensions of irreducible ...}{}
  \set:bib'journal{Fund. Math.}{}
  \set:bib'date{1962/63}{}
}

The extra empty arguments on each line are for auxiliary properties (see below). What happens thereafter with \rsk@toks depends on the code in the last arg of \RestrictedSetKeys.

\section{Auxiliary properties}

Unfortunately, the previous section isn’t the entire story. In addition to the values of each field, we need to store a set of auxiliary properties associated with those values. Note that properties are explicitly associated with values, not with keys, because each value of an additive key could have different properties.

All such extra data will be stored in a special field named “aux”, with embedded tags to indicate which field each piece of the field is associated with. The extra bits can be extracted on demand using standard techniques, and the primary value of each field is not burdened with any attachments, so that comparisons or scanning of the field contents can remain as simple as possible.
Thus in practice there is at least one bit of auxiliary information in every bib item, and our previous example would have the title language indicated:

\DSK@def\bib'title{Eine Bemerkung zur Darstellung von Polynomen
"{u}ber Verb""{a}nden}
@append\bib'title\bib'aux{\selectlanguage{german}}

\set@property
23 \def\set@property#1{%
24 \begingroup
25 \edef\@tempa{\extract@group#1}%
26 \edef\@tempa{
27 \@nx@append\@nx#1\@xp\@nx\csname \@tempa,aux\endcsname
28 }%
29 \@xp@endgroup
30 \@tempa
31 }

\get@property
32 \% \get@property\destination\bib'title
33 \def\get@property#1#2{%
34 \get@nth@property#1#2\m@ne
35 }

\get@nth@property
36 \% \get@nth@property\destination\bib'title N
37 \def\get@nth@property#1#2#3{%
38 \begingroup
39 \edef\@tempa{\extract@group#2}%
40 \@tempcnta#3\relax
41 \@tempcntb\z@
42 \@xp\scan@properties\@xp#2\csname \@tempa,aux\endcsname
43 \edef\@tempa{\def\@nx#1{\@tempa}}%
44 \@xp@endgroup
45 \@tempa
46 }

\scan@properties
47 \def\scan@properties#1#2{%
48 \begingroup
49 \def\@tempa{#1}%
50 \let\@tempc\@empty
51 \@xp\find@property #2 \@nil\@nil
52 \edef\@tempa{\def\@nx#1\@tempa{\@tempc}}%
53 \@xp@endgroup
54 \@tempa
55 }

\find@property
56 \def\find@property#1#2{%
\ifx\@nil#1\%
\else
  \def\@tempb{#1}\%
  \ifnum\@tempcnta<\z@
    \def\@tempc{#2}\%
  \else
    \advance\@tempcntb\@ne
    \ifnum\@tempcntb=\@tempcnta
      \def\@tempc{#2}\%
    \fi
  \fi
  \fi
\fi
\@xp\find@property
\fi
}

\reset@property
\def\reset@property#1#2{%}
\reset@nth@property#1\m@ne{#2}%
\eset@nth@property
% \reset@nth@property\bib\'title N VALUE
\def\reset@nth@property#1#2#3{%
  \begingroup
  \edef\@tempa{\extract@group#1}\%
  \@tempcnta#2\relax
  \@temptokena{#3}\%
  \toks@\emptytoks
  \@tempcntb\z@\%
  \@xp\reset@scan\@xp#1\csname\@tempa,aux\endcsname
  \edef\@tempa{%
    \def\@xp\@nx\csname\@tempa,aux\endcsname{\the\toks@}\%
  }\%
  \@xp\endgroup
  \@tempa
\}

\reset@scan
\def\reset@scan#1#2{%
  \begingroup
    \def\@tempa{#1}\%
    \@xp\reset@scan0a\#2\\@nil\@nil
    \edef\@tempa{\toks@{\the\toks@}}\%
    \@xp\endgroup
    \@tempa
\}
\reset@tempa

Some machinery for finite state automata

Coincidentally I needed to write two finite state automaton parsers for two related packages, so after writing them separately I spent some time analyzing the code fragments they shared in common and abstracted them so that the cs names could be shared.

\fssa{\textbf{FSA lookahead.}}
\def\fssa{(\futurelet\@let@token\fssa{t})}

\fssb{\textbf{FSA bypass a token.} Don’t delete the space at the end!}
\def\fssb{(\afterassignment\fssa\let\@let@token= )}

\fssc{\textbf{FSA copy a token} (not space, bgroup, egroup).
\def\fssc{(\aftergroup\fssa\let\@empty)\fssa{t}}}

\fssn{\textbf{FSA next action.} This is just a placeholder definition.}
\let\fssn\@empty

\fss{\textbf{FSA test.} This is just a placeholder definition.}
\let\fss\@empty
6 Now some of the real work

\newtoks\rsk@toks
\newtoks\rsk@toks

\kvIfEmpty Beginning here.
\def\kvIfEmpty#1#2{% 
\expandafter\ifx\csname#1'#2\endcsname\empty 
\@xp\@firstoftwo 
\else 
\@xp\@secondoftwo 
\fi 
}

\kvIfAdditive
\def\kvIfAdditive#1{% 
\let\@let@token \csname \kv@setter#1\endcsname 
\afterassignment\@nilgobble 
\let\@let@token \@let@token \@empty\@empty\@nil 
\ifx\@let@token\DSK@append 
\@xp\@firstoftwo 
\else 
\@xp\@secondoftwo 
\fi 
}

\kv@setter It irritates me that I can't embed the \csname and \endcsname in here.
\def\kv@setter#1{set:\@xp\@gobble\string#1}

\kv@DSAK Define a simple, superseding, or additive key.
\def\kv@DSAK#1#2{% 
\addto@group@reset#1{\let#1\@empty} 
\edef\@tempa{\def\csname \kv@setter#1\endcsname} 
\@tempa{#2#1} 
}

\kv@DDK This function is used for a dummy key whose value (expansion) should be empty 
but that should appear non-empty to \kvIfEmpty.
\def\kv@DDK#1{% 
\addto@group@reset#1{\def#1{\@empty}}% 
\edef\@tempa{\def\csname \kv@setter#1\endcsname}% 
\@tempa{#2#1} 
}

\DSK@def
\edef\DSK@def#1{% 
\ifx#1\@empty\else 
\PackageWarningNoLine{rkeyval}{Key \string#1 should not be repeated}% 
\fi 
}
We weed out empty field values for consistency with \DSK@append.

```latex
\def\DSK@redef#1#2{%
  \ifempty{#2}{\@gobble}{%
    \def#1{#2}%
    \set@property#1%
  }%
}%
```

```latex
\def\init@group@reset#1{%
  \begingroup
  \edef\@tempb{\@xp\@nx\csname #1@reset\endcsname}%
  \@xp\ifx\@tempb\relax
    \@xp\xdef\@tempb{\let \csname #1,aux\endcsname\@nx\@empty}
  \fi
  \endgroup
}%
```

```latex
\def\addto@group@reset#1{%
  \begingroup
    \edef\@tempa{\extract@group#1}%
    \init@group@reset\@tempa
    \edef\@tempa{\@nx\g@addto@macro\@xp\@nx\csname\@tempa @reset\endcsname}%
  \@xp\endgroup
  \@tempa
}%
```

```latex
\newcommand{\DefineSimpleKey}[2]{%
  \@[rkv\@DSAK\csname #1'#2\endcsname}{\DSK@def}%
}%
```

```latex
\newcommand{\DefineSupersedingKey}[2]{%
  \@[rkv\@DSAK\csname #1'#2\endcsname}{\DSK@redef}%
}%
```

```latex
\newcommand{\DefineAdditiveKey}[3]{%
  \@[rkv\@DSAK\csname #1'#2\endcsname}{\DSK@add}%
}%
```
We weed out empty field values (e.g., editor=\{} or editor=\{}\) because otherwise an additive field could end up with a value like \\name{} which appears non-empty to \rkvIfEmpty but produces no output on the page.

\def\DSK@append#1#2#3{\@ifempty{#3}{\@gobble}{\@append#1#2{#3}\set@property#2}}

\DefineDummyKey
\newcommand{\DefineDummyKey}[2]{\@xp\rkv@DDK \csname #1'#2\endcsname}

\RestrictedSetKeys
\newcommand{\RestrictedSetKeys}[3]{\global\rsk@toks\@xp{\csname #2@reset\endcsname}\def\rsk@finish{#3}\gdef\rsk@set{\@xp\rsk@set@a\csname#2'}#1\relax\begingroup\rsk@changecase\aftergroup\rsk@set\lowercase{\rsk@set FIELDNAME\endcsname}---\@xp\rsk@set@a\csname bib'fieldname\endcsname---\rsk@set@a\bib'abcdef

\rsk@unknown@key
\def\rsk@unknown@key#1{\PackageWarning{rkeyval}{Unknown key: \string#1}\@xp\def\csname\rkv@setter#1\endcsname{\DSK@redef#1}}
7 The state machine

State 0: Skip opening brace (\rsk@z).
  space -> 0
  { -> 2
  other -> error "Missing open brace"

State 1: Skip comma (\rsk@a).
  space -> 1
  \par -> 1
  comma -> 2
  @ -> read optional arg; 1
  } -> 6
  other -> error "Missing comma"; 2

State 2: Find field name (\rsk@b).
  space -> 2
  \par -> 2
  comma -> 2
  letter -> 3
  { -> error "Missing key name"; 4
  } -> 6
  other -> error "Invalid key name character"; 2

State 3: Scan field name (\rsk@c).
  letter -> 3
  comma -> error "Invalid key name character"; 3
  equal -> 4
  other punct -> 3
  space -> 4
  { -> error "Missing equal sign"; 4
  } -> error "Missing equal sign"; 4
  other -> error "Invalid key name character"; 3

State 4: Skip equals (\rsk@d).
  space -> 4
  equal -> 4
  { -> 5
  other -> error "Missing \{ for value of current key"; 5

State 5: Read field value (\rsk@set@a).
  any -> 1

State 6: Done (\rsk@end).

\rsk@z State 0: Skip opening brace.

\def\rsk@z{%
  \ifx\bgroup\@let@token
  \let\fsa@t\rsk@b
  \let\fsa@n\fsa@b
  \else
  \fi
\let\fsa@n\fsa@b \else \fi \fi \fi \fsa@n
\rsk@a

State 1: Skip comma.
\def\rsk@a{% 
\ifx\@let@token\@sptoken \let\fsa@t\rsk@b \let\fsa@n\fsa@b \else \fi \fi \else \ifx,\@let@token \endgroup \let\fsa@t\rsk@b \let\fsa@n\fsa@b \else \ifx\egroup\@let@token \endgroup \let\fsa@n\rsk@end \else \endgroup \let\fsa@n\rsk@erraa \fi \fi \fi \fsa@n
\rsk@b

State 2: Find field name.
\allowpar here to permit a blank line after the end of one key-val pair and the start of the next (perhaps to break up a long list into sections).
\def\rsk@b{% 
\ifcat\@nx\@let@token A% \let\fsa@t\rsk@c \let\fsa@n\fsa@b \else \ifx\@sptoken\@let@token \let\fsa@t\rsk@c \let\fsa@n\fsa@b \else \fi \fi \else \ifx\@sptoken\@let@token \let\fsa@t\rsk@c \let\fsa@n\fsa@b \else \fi \fi \rsk@bb \fi \fi \fi \fsa@n
}
\rsk@bb
\def\rsk@bb{%
275 \ifex,{\@let@token
276 \let\fsa@n\fsa@b
277 \else
278 \ifx\bgroup\@let@token
279 \let\fsa@n\rsk@errb
280 \else
281 \ifx\egroup\@let@token
282 \let\fsa@n\rsk@end
283 \else
284 \ifx\par\@let@token
285 \let\fsa@n\rsk@errc
286 \else
287 \let\fsa@n\rsk@errc
288 \fi
289 \fi
290 \fi
291 \fi
292 \fi
293 }
\rsk@c State 3: Scan field name.
\def\rsk@c{%
294 \ifcat\@nx\@let@token A
295 \let\fsa@n\fsa@c
296 \else
297 \ifx@sptoken\@let@token
298 \let\fsa@t\rsk@d
299 \let\fsa@n\fsa@b
300 \else
301 \ifx=\@let@token
302 \let@saw@equal T
303 \let\fsa@t\rsk@d
304 \let\fsa@n\fsa@b
305 \else
306 \rsk@cb
307 \fi
308 \fi
309 \fi
310 \fi
311 \fsa@n
312 }
\rsk@cb
\def\rsk@cb{%
313 \ife\@let@token

\let\fsa@n\rsk@errc
\else
  \ifcat\@nx\@let@token .\%
    \let\fsa@n\fsa@c
  \else
    \ifx\bgroup\@let@token
    \let\fsa@n\rsk@noequal
  \else
    \ifx\egroup\@let@token
    \let\fsa@n\rsk@noequal
  \else
    \let\fsa@n\rsk@errc
  \fi
  \fi
\fi
\fi
}
\saw@equal
\let\saw@equal=F
\rsk@d State 4: Skip equals.

  If no equal sign ever came along, then give a warning about it and set \saw@equal to true so that when \rsk@noequal cycles through again it will take the other branch.
\def\rsk@d{%\n  \ifx\bgroup\@let@token
    \ifx\saw@equal T\%
      \aftergroup\endcsname
      \let\fsa@n\rsk@endcase
    \else
      \let\fsa@n\rsk@noequal
      \aftergroup\endcsname
    \fi
  \else
    \ifx\egroup\@let@token
      \let\fsa@n\rsk@noequal
    \else
      \let\fsa@n\rsk@errc
    \fi
  \fi
\else
  \ifx\@sptoken\@let@token
    \let\fsa@n\fsa@b
  \else
    \ifx=\@let@token
      \let\saw@equal T\%
      \let\fsa@n\fsa@b
    \else
      \let\fsa@n\rsk@erre
    \fi
  \fi
\fi
\fi
\fi
\fsa@n}
Here we get improved reporting of error context by changing end-of-line to be different from normal space. If we don’t find a comma on the current line, assume it is an error.

State 5: Read field value.

State 0b
\rsk@end  
State 6: Done. 
Lets see, now why did I add this?
\rsk@end  
\def\rsk@end{% 
  \global\let\rsk@set\rsk@terminate 
  \rsk@endcase 
  \endgroup 
  \endcsname 
  \afterassignment\rsk@finish 
  \toks\@group 
}

\rsk@terminate  
\def\rsk@terminate{%\exp@gobble\csname}

\NoCommaWarning  
\def\NoCommaWarning{\PackageWarning{rkeyval}{Missing comma}}%

\NoCommaError  
\% \def\NoCommaError{\rsk@err{Missing comma}\@ehc} 
\% 
\rsk@nocomma  
\def\rsk@nocomma{\NoCommaWarning}

\rsk@err  
\def\rsk@err{\PackageError{rkeyval}}

\rsk@errf  
\def\rsk@errf{\rsk@err{Missing open brace}\@ehc\rsk@b}

\rsk@erraa  
\long\def\rsk@erraa{\rsk@nocomma \let\fsa@t\rsk@t \fsa@l}

\rsk@errb  
\def\rsk@errb{\rsk@err{Missing key name}\@ehc\rsk@d}

\rsk@errc  
\def\rsk@errc{\rsk@err{Invalid key name character}\@ehc\fsa@b}

\rsk@noequal  
\def\rsk@noequal{\rsk@err{Missing equal sign}\@ehc\rsk@d}
In this case we guess that the value is without braces but probably terminated with a comma. We want to scan up to the comma in order to get back in synch.

\begin{verbatim}
413 \def\rsk@erre#1,\%
414 \rsk@err{Missing open brace for key value}\@ehc
415 \iffalse\fi
416 \endgroup
417 \endcsname
418 \rsk@endcase }{#1},\%
419 }
\end{verbatim}

\texttt{PopCatcodes}

The usual \texttt{endinput} to ensure that random garbage at the end of the file doesn’t get copied by \texttt{docstrip}.

\texttt{endinput}

\textbf{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; numbers in roman refer to the code lines where the entry is used.

\begin{center}
\begin{tabular}{llll}
\textbf{Symbols} & \textbf{\@append} & \textbf{\@emptytoks} & \textbf{\@firstoftwo} \\
& \textbf{\@gobblethree} & \textbf{\@ifempty} & \textbf{\@ifundefined} \\
\textbf{\@ifdef} & \textbf{\@gobble} & \textbf{\@ifundefined} & \textbf{\@nilgobble} \\
\textbf{\@popcatcodes} & \textbf{\@secondoftwo} & \textbf{\@secondoftwo} & \textbf{\@sptoken} \\
\textbf{\@temptokenb} & \textbf{\@xp} & \textbf{\~} & \textbf{\@ccode} \\
\textbf{\@addto@macro} & \textbf{\@addto@group@reset} & \textbf{\@afterassignment} & \textbf{\@aftergroup} \\
\textbf{\@group} & \textbf{\@group} & \textbf{\@group} & \textbf{\@group} \\
\textbf{\@getline@char} & \textbf{\@extract@group} & \textbf{\@extract@group@a} & \textbf{\@extract@group@b} \\
\textbf{\@extract@group@c} & \textbf{\@extract@group@d} & \textbf{\@extract@group@e} & \textbf{\@extract@group@f} \\
\textbf{\@extract@group@g} & \textbf{\@extract@group@h} & \textbf{\@extract@group@i} & \textbf{\@extract@group@j} \\
\textbf{\@extract@group@k} & \textbf{\@extract@group@l} & \textbf{\@extract@group@m} & \textbf{\@extract@group@n} \\
\textbf{\@extract@group@o} & \textbf{\@extract@group@p} & \textbf{\@extract@group@q} & \textbf{\@extract@group@r} \\
\textbf{\@extract@group@s} & \textbf{\@extract@group@t} & \textbf{\@extract@group@u} & \textbf{\@extract@group@v} \\
\textbf{\@extract@group@w} & \textbf{\@extract@group@x} & \textbf{\@extract@group@y} & \textbf{\@extract@group@z} \\
\textbf{\@find@property} & \textbf{\@lccode} & \textbf{\@lccode} & \textbf{\@lccode} \\
\end{tabular}
\end{center}