dowith.sty

—

Apply Command to Each Item in a List of Arguments in “\TeX’s Gullet”*

Uwe Lück†

November 27, 2012

Abstract

This package provides macros for applying a “command” to all items in a “list of possible macro arguments,” and also for extending and reducing macros storing such lists. “Brace groups” are single items of such lists, as opposed to token lists. Iteration is implemented within \TeX’s expansion processor, so works within \texttt{\write} as with blog.sty. Loop and list macros in other packages are discussed in the documentation. There is no need for \TeX{} to which some of them refer.

The package is “generic,” i.e., should also work with Plain \TeX{} or even other formats, relying on the plainpkg package for some minimal \LaTeX{}-like behaviour.

Related packages: catoptions etextools etoolbox forarray forloop loops |multido moredefs lmake texapi xfor xspace

Keywords: programming structures; macro programming, loops, list macros

Contents

1 Usage and Features 2
1.1 Installing and Calling ........................................... 2
1.2 What It Does With What Lists ................................... 3
1.3 The Notion of Arglists for \LaTeX{} Users .......................... 3
1.4 Anatomy of \TeX{} .................................................. 4
1.5 \TeX{}’s Tokens .................................................... 5
1.6 Arglists vs. Lists of Tokens—Example .......................... 6
1.7 Another Notation and the Example’s Steps ...................... 7

*This document describes version v0.3 of dowith.sty as of 2012/11/05.
†http://contact-ednotes.sty.de.vu
1 Usage and Features

1.1 Installing and Calling

The file dowith.sty is provided ready, installation only requires putting it somewhere where \TeX{} finds it (which may need updating the filename database)\footnote{http://www.tex.ac.uk/cgi-bin/texfaq2html?label=inst-wlcf}. The packages \texttt{plainpkg}\footnote{http://ctan.org/pkg/plainpkg} and \texttt{stacklet (catcodes)}\footnote{http://ctan.org/pkg/catcodes} must be installed as well.

As to calling (loading): dowith is a “\texttt{plainpkg} package” in the sense of the \texttt{plainpkg} documentation that you may consult for details. So roughly,

\begin{itemize}
  \item load it by \texttt{\usepackage{dowith}} if you can,
  \item otherwise by \texttt{\RequirePackage{dowith}} (perhaps from within another “\texttt{plainpkg} package”),
  \item or by \texttt{\input{dowith.sty}}
\end{itemize}
1.2 What It Does With What Lists

The single commands that the package provides are described in the implementation section below. What follows here is some general background about how the commands work.

The term ‘list’ may refer to various things and need clarification here.

First of all, we are not referring to \texttt{\LaTeX} \texttt{list} environments such as \texttt{enumerate} or \texttt{itemize}; neither to “TODO” lists of what needs to be done soon.

Rather, \texttt{dowith} allows you to abbreviate

\[
\langle \text{cmd} \rangle \langle \text{arg-1} \rangle \langle \text{cmd} \rangle \langle \text{arg-2} \rangle \ldots \langle \text{cmd} \rangle \langle \text{arg-n} \rangle
\]

by

\[
\texttt{\DoWith}(\text{cmd})\langle \text{arg-1} \rangle \langle \text{arg-2} \rangle \ldots \langle \text{arg-n} \rangle \texttt{\StopDoing}
\]

or by

\[
\texttt{\DoWithAllOf}(\text{cmd})\{\langle \text{arg-1} \rangle \langle \text{arg-2} \rangle \ldots \langle \text{arg-n} \rangle\}
\]

With small \(n\), one may doubt whether this really is an abbreviation \ldots; anyway,

\[
\langle \text{arg-1} \rangle \langle \text{arg-2} \rangle \ldots \langle \text{arg-n} \rangle
\]

was an attempt to refer to the kind of lists we are dealing with.

\[
\langle \text{arg-1}, \text{arg-2}, \ldots, \text{arg-n} \rangle
\]

are the “items” of the list. The question is: what counts as an item?

We might say that \texttt{aa} is a list of \textit{two} items, \langle \text{arg-1} \rangle being \texttt{a} and \langle \text{arg-2} \rangle being \texttt{a}, too.

When we do \textit{three} keystrokes to get \texttt{a\_a} instead of \texttt{aa}, we still have \textit{two} items, \langle \text{arg-1} \rangle being \texttt{a} and \langle \text{arg-2} \rangle being \texttt{a} too. Strange, isn’t it?

Also, when in \texttt{aa} we replace the first \texttt{a} by a backslash, \texttt{\}, we get \texttt{\_a}, and this is a list of a \textit{single} item, \langle \text{arg-1} \rangle = \texttt{\_a} \ldots

You shouldn’t believe these stories of mine entirely. What I am alluding to is that the “items” \texttt{dowith} is about are determined in terms of \texttt{\TeXX}’s \textit{tokens}, and the relation between the “characters you type” and \texttt{\TeXX}’s \textit{tokens} is not entirely straightforward.

1.3 The Notion of Arglists for \LaTeXX Users

Still, it may suffice to clarify what counts as an \langle \text{arg-i} \rangle without speaking of \textit{tokens} explicitly: It is simply what a \textit{one-parameter macro} (where the parameter is \textit{not delimited} in terms of \texttt{\TeXXbook} pp. 203f.) can take as an \textit{argument}.

The lists \texttt{dowith} is about then are lists of \textit{possible arguments} in the previous sense—let me call them “arglists.”\footnote{Not to be confused with German \texttt{Arglist}.} The single \textit{items} of such lists are those single
possible arguments. They become actual arguments beginning from the leftmost possible one when \texttt{dowith} presents them to that \langle cmd \rangle mentioned earlier—where \langle cmd \rangle should be a one-parameter macro (or some \TeX\ primitive parsing arguments similarly).

The reader perhaps has an intuitive understanding of what can be an argument of a one-parameter macro. A strict \LaTeX\ user may think that such an argument \langle arg-i \rangle just has form \{\langle ark-i \rangle\}, i.e., \langle arg-i \rangle = \{\langle ark-i \rangle\} for some \langle ark-i \rangle. Such arguments are also called “brace groups”. (\LaTeX\’s optional arguments \{\langle extra \rangle\} do not count as possible arguments here, they are not macro arguments in the sense of \TeX\book.) In this restricted \LaTeX\ sense, arglists consist of brace groups

\{\langle ark-1 \rangle\}\{\langle ark-2 \rangle\} \ldots \{\langle ark-n \rangle\},

and each single brace group is an item of it.

The \TeX\ macro writer, by contrast, knows that a macro argument doesn’t need outer braces. In an intuitive sense, a single “command” can be a macro argument, too. “Command” may be understood as “control sequence” (starting with a backslash), but some authors also have considered single characters (character tokens?) “commands.” Blank spaces, by contrast, are ignored when a macro looks for its argument.

1.4 Anatomy of \TeX

The documentation of v0.22 as of 2012-06-04 said that the package is about “lists in \TeX\’s mouth.” However, this was very wrong. I believed it following Alan Jeffrey’s paper “Lists in \TeX\’s Mouth”\(^5\) in whose Section 2 you read:

\TeX\’s programming facilities come in two forms—there are \TeX\’s macros which are expanded in its mouth, and some additional assignment operations like \texttt{\textbackslash def} which take place in the stomach.

The macros that Jeffrey lists and describes in that article can be obtained as a \CTAN\ package \texttt{lambda-lists}\(^6\). If you follow the link given here (in the footnote), you currently (2012-11-03) read about this package:

These list-processing macros avoid the reassignments employed in the macros shown in Appendix D of the TeXbook: all the manipulations take place in what Knuth is pleased to call “\TeX\’s mouth”.

But Knuth doesn’t. On page 267 of \TeX\book, you read:

Chapter 7 has described the process by which input files are converted to lists of tokens in \TeX\’s “mouth,” and Chapter 20 explained how expandable tokens are converted to unexpansible ones in \TeX\’s “gullet” by a process similar to regurgitation.

\(^6\)http://ctan.org/pkg/lambda-lists
I.e., the “mouth” is \TeX’s “tokenizer,” the inner part of what van Eijkhout calls \TeX’s “input processor” on, e.g., p. 15 of his \TeX by Topic.\footnote{It is available as a CTAN package texbytopic at http://ctan.org/pkg/texbytopic} The exact rules the tokenizer follows are described on pp. 46f. (Chapter 8!) of The \TeXbook. Macro expansion takes place in \TeX’s “gullet”, which van Eijkhout calls \TeX’s “expansion processor” (p. 16). Abrahams, Hargreaves, and Berry follow Knuth’s terminology on pp. 16 and 46f. of their \TeX for the Impatient.\footnote{It is available as CTAN package impatient, http://ctan.org/pkg/impatient}

\TeX’s gullet has been called “\TeX’s mouth” also in the documentation of my \textsc{bitelist}\footnote{http://ctan.org/pkg/bitelist} package and in the documentation of the package \texttt{bibleref-mouth}\footnote{http://ctan.org/pkg/bibleref-mouth}.

Moreover, I should have clarified that Jeffrey’s paper deals with “lists” in some general, rather abstract sense, different from the kind of lists the present documentation tries to characterize as the objects for \texttt{dowith}.

1.5 \TeX’s Tokens

The \texttt{dowith} package is a tool that affects the order of tokens in \TeX’s gullet.

The “characters you type” enter “\TeX’s mouth” line by line, in a slightly modified appearance. Each line forms a \textit{string}. \TeX takes initial substrings away from it and turns them into \textit{tokens} that are appended to the right of \TeX’s \textit{expansion buffer} (“gullet”).

There are two kinds of tokens here: named tokens and character tokens. “Named” tokens usually are referred to as “control sequence tokens” or just “control sequences”—I really want to avoid those horrible confusions from The \TeXbook. There never are any “parameter tokens” in \TeX’s gullet (perhaps unless one considers a one-step macro expansion a two-or-more-step procedure). The character(s) after the escape character until some delimiting character form a \textit{string} that is the \textit{name} of the token that is formed—a named token, as I am saying. Character tokens are formed by removing a character from the beginning of the character buffer and appending it to the token buffer paired with its \textit{category} code.

For every \textit{string} of characters, there is exactly one (possible) \textit{named token} whose name the string is\footnote{“Possible” refers to the fact that \TeX does not store named tokens anywhere before they appear in its gullet, maybe apart from “primitive” tokens that have a “pre-assigned meaning” when a \TeX run starts.—What is more bad with my claim is that the \TeX program by design cannot extend its memory arbitrarily—even not using the “cloud”—, so it doesn’t support tokens whose name lengths are above a certain limit.}. It is so common (starting from The \TeXbook) to denote the token whose name \langle string \rangle by ‘\textbackslash \langle string \rangle’. For instance, the token whose name input is denoted by ‘\textbackslash input’. On the other hand, on page 7 of The \TeXbook ‘\textbackslash input’ is a “string of characters.” With this notation, it is already difficult to explain what the L\LaTeX\ command \texttt{\DeclareRobustCommand} does or what the difference between a starred \LaTeX command and a starred \LaTeX environment is.\footnote{A reader knowing \LaTeX only thinks that ‘\textbackslash\textbackslash’ is the result of typing a double backslash and a space and that ‘\textbackslash equation\*’ is the “command” \texttt{\equation} followed by a \texttt{*}.} The \TeXbook makes it worse by saying on page 39: “A control sequence is considered to be a single object that is no longer composed of a sequence of symbols.” So “it depends” whether ‘input’ is a string of characters or not—it is \textit{before} tokenization, but no longer afterwards. So if you have two computers and start a \TeX run on each of them with a little difference in time, there will be a moment where ‘input’ is a string on the one computer but not on the other? This appears to me like saying “When we apply the square root function to the number 4, the number 4 will no longer be the number 4, it will be the number 2 instead.”

\footnote{It is available as a CTAN package texbytopic at http://ctan.org/pkg/texbytopic}
The \TeX{}book does offer an alternative notation for named tokens: “boxing;” so the token whose name is `\texttt{input}` can be denoted by the rather “graphical” notation `\texttt{ntok(input)}` (used only exceptionally). I would suggest something like `\texttt{ntok(input)}` for clarity and `\texttt{\textbf{?input}}` for brevity.

Named tokens may get into \TeX{}’s gullet by “tokenization” as described above, i.e., they are drawn from the character buffer. But they also can appear in \TeX{}’s gullet “from within,” by the manipulation inside \TeX{}’s gullet.

More formally, those manipulations are called “expansion,” and \TeX{}’s gullet can be conceived of as a token buffer that is feeded to the right (or end) by tokenization from the character buffer. Expansion means that certain tokens in the token buffer are substituted by other ones. This way tokens may get into \TeX{}’s gullet that emerged from tokenization a “long time ago”, maybe in a previous run that created the format (\TeX{}’s variant \texttt{\textsc{Initex}}); or tokens may appear by some hardwired expansion function.

However, named tokens may get into \TeX{}’s gullet also by expansion, never having been drawn by tokenization and not being hardwired. This happens by the `\texttt{\texttt{\textbf{\textbackslash csname}}}` construct. The input code may contain

\begin{verbatim}
\texttt{\textbackslash csname\textunderscore tupni}endcsname
\end{verbatim}

This may be converted into 7 tokens entering \TeX{}’s gullet, the first one being `\texttt{ntok(csname)}`, the last one `\texttt{ntok(endcsname)}`, and five character tokens in between. Due to some function (which I would denote as `\texttt{\textbf{\textbackslash csname}}`) originally associated with the token `\texttt{ntok(csname)}`, those seven tokens then are replaced by `\texttt{ntok(tupni)}`, the named token whose name is `\texttt{tupni}`. It is not required that the \TeX{} program knows about a token `\texttt{ntok(tupni)}`, neither anybody must type `\texttt{\textbackslash tupni}` in any file.

1.6 Arglists vs. Lists of Tokens—Example

Let us reconsider the examples from Sections 1.2 and 1.3, and pack them into a single example. If you type a file line

\begin{verbatim}
a\textbackslash a\{a\}
\end{verbatim}

(eight keystrokes), it should usually be converted into this seven-item list of (five) tokens:

\begin{verbatim}
a_11 a_10 a_{11} ntok(a) \{ a_{11} \}_2
\end{verbatim}

—with notation from Section 1.5 and \TeX{}book’s notation `\texttt{\textbackslash \langle char\rangle_{\langle cat\rangle}}` for the character token that \TeX{}’s tokenizer forms from `\texttt{\langle char\rangle}` in the character buffer when `\texttt{\langle char\rangle}`’s category code is `\texttt{\langle cat\rangle}`.

\begin{itemize}
\item \texttt{\textbackslash csname\textunderscore tupni}endcsname
\end{itemize}

\begin{itemize}
\item The box notation is introduced on page 38 without explanation, as if it explained something.
\item I am suggesting the question mark for named tokens since \TeX{} “must look up the current definition” of a named token according to \TeX{}book p. 39, while the meaning of character tokens rather is “fixed,” at least according to \TeX{}book p. 39. However, active-character such as \texttt{\textbackslash .} are in the same situation as named tokens as to this respect. The dot notation may be fine for them, though.
\item These considerations may not be essential here, rather a draft for a paper. Using \texttt{dowith}, one better just thinks of the arglist items one actually lists.
\end{itemize}
It turns out that the token list in 2 provides an arglist of four items: The token \texttt{a} at the first and third place, the named token \texttt{ntok(a)}, and the entire token list \texttt{\{a\}2} as a single item—a “brace group.” The space token is ignored.\footnote{The \texttt{T\!\TeX}book p. 201: “\texttt{T\!\TeX} doesn’t use single spaces as undelimited arguments.”}

You can try this after \texttt{\renewcommand{\a}{A}}\footnote{Otherwise \texttt{\a} is a one-parameter macro that breaks \texttt{dowith}’s control.} with \texttt{dowith}:

\begin{verbatim}
\DoWith\typein{a\a\a\{a\}}\StopDoing
\end{verbatim}

Then \texttt{T\!\TeX} shows \texttt{a, a, A} from \texttt{\a}, and another \texttt{a} from within the braces—\texttt{\typein} (as any macro with arguments) removes them.

I have avoided saying 2 were an arglist of 4 items. The mathematical basic way of writing lists—understood as finite sequences—as “comma-separated lists” within brackets may clarify the difference (that the juxtaposition notation tends to conceal). The \texttt{token} list is

\begin{verbatim}
\{a_{11}, \texttt{\{\}, a_{11}, \texttt{ntok(a)}, \{1, a_{11}\}_2}\}
\end{verbatim}

while the list of macro arguments is

\begin{verbatim}
\{a_{11}, a_{11}, \texttt{ntok(a)}, \{1, a_{11}, \}_2\}.
\end{verbatim}


\begin{verbatim}
2 or 4 simply is not an arglist (since neither \{1 nor \}_2 can be a macro argument), and the arglist 5 “provided” by the list of tokens is not a list of tokens—its final item is a three-item list of tokens, and a token cannot be a list of two or more tokens itself(?).
\end{verbatim}

1.7 Another Notation and the Example’s Steps

To write token lists easier and hopefully easier to read, I would suggest writing ‘\langle char\rangle’ for the character token that the tokenizer “usually” forms from character ⟨char⟩, i.e., adding the standard category code as in The \texttt{T\!\TeX}book (page 37). Then 2 would read\footnote{See Section 1.5 for the question mark.}

\begin{verbatim}
.a_{11} a_{11} \texttt{ntok(a)} \{1, a_{11}\}_2
\end{verbatim}

and the corresponding arglist is

\begin{verbatim}
.a_{11} a_{11} \texttt{ntok(a)} \{1, a_{11}\}_2
\end{verbatim}

In “retrospect,” the result of tokenizing 2 should be

\begin{verbatim}
?\DoWith?\typein{a\a\a\{a\}?\StopDoing
\end{verbatim}

and the intention is that it works like

\begin{verbatim}
?\typein{\a}\typein{\a}\typein{\a}\typein{\texttt{ntok(a)}}?\StopDoing
\end{verbatim}

(The definition of \texttt{\DoWith} in Section 3.3.1 indeed adds surrounding braces, if missing.) However, \texttt{T\!\TeX} rather tries to work with as few tokens ahead as possible. When it finds ?\texttt{DoWith} and the latter’s meaning is the one intended by \texttt{dowith}, it first looks for nothing more than
1 USAGE AND FEATURES

the two arguments required by our definition of DoWith. A few moments later, the token buffer’s content will just be:

\begin{verbatim}
\end{verbatim} (10)

Next ?typein{.a} is expanded according to the code for typein in latex.ltx. Some unpoppable tokens will emerge and be moved into the “instruction buffer,” and you should get a screen message with a and a prompt. When you have entered something, the remaining expandafter tokens and the ?fi will be removed from the character buffer, and it contains only

\begin{verbatim}
?DoWith?typein
\end{verbatim} (11)

Another token is ordered from the tokenizer to provide a second argument for expanding DoWith. The token .\ is in, but that doesn’t serve as a macro argument. It is removed, and the next token is .a. The same story as before happens, until the named token ?a is found...

1.8 Summary of Possible Arglist Items

For $0 \leq i \leq 15$, let $X_i$ be the set of character tokens of category code $i$. $X_1$ is the set of tokens working like \{1, and $X_2$ is the set of tokens working like \).

Let $E$ be the set \{3, 4, 6, 7, 8, 11, 12, 13\}. These numbers are the category codes for math, align, parameter, super, sub, letter, other, active respectively. Let $X_E$ be the set of character tokens of category code in $E$ (so $X_E = \bigcup_{i \in E} X_i$).

Let $\circ$ be the concatenation operation among token lists. The following kinds of token lists form a single arglist item, i.e., can serve as an argument for an undelimited parameter:

1. a named token, or the single-token list consisting of it, if you prefer that;
2. a character token from $X_E$ or the list consisting of it;
3. a brace group. That is a token list meeting the following conditions: (i) its first token is in $X_1$, (ii) its last token is in $X_2$, (iii) it has as many occurrences of tokens from $X_1$ as from $X_2$, (iv) if it is split as $\lambda \circ \rho$, there are not more $X_2$ occurrences in $\lambda$ than $X_1$ occurrences in $\rho$ (“don’t close before opening”).

The second claim can be checked with \DoWith\typein$#^{a_1}$\StopDoing (12)
as to what works. (The claim is not affected by one or two surprises.)\footnote{If you use \DoWithAllOf\typein{a_a\{a(a)}} instead, the entire token sequence \texttt{8} will appear in the token buffer “at once.”} Characters with different category codes either are not converted into a character token or are not accepted as macro arguments. The latter applies to “brace” tokens in $X_1$, $X_2$ and to the single space token \texttt{.10}.

As to “brace groups”, the third and fourth condition above are intended to say that what is between the two outer tokens is \langle balanced text \rangle in the sense of The \textit{TeX}book pp. 275f. and 385; i.e., for two tokens $a$, $b$ and a token list $\beta$, $(a) \circ \beta \circ (b)$ is a brace group exactly if $a$ is from $X_1$, $b$ is from $X_2$, and $\beta$ is \langle balanced text \rangle. The conditions are more formal than what I can find in The \textit{TeX}book, but still they don’t give me an idea of all possibilities. This should be improved by the following recursive definition:

\begin{verbatim}
\DoWith\typein$#^{.a_1}$\StopDoing
\end{verbatim} as to what works. (The claim is not affected by one or two surprises.)\footnote{TODO: Define for representations by maps, or: “Concatenation is about as basic as natural numbers and is understood in terms of axioms rather than by a definition.”—See notes from 2011 (even with attempts with Category theory) the English Wikipedia for sequences—German article too much restricted to maps.} Characters with different category codes either are not converted into a character token or are not accepted as macro arguments. The latter applies to “brace” tokens in $X_1$, $X_2$ and to the single space token \texttt{.10}.

As to “brace groups”, the third and fourth condition above are intended to say that what is between the two outer tokens is \langle balanced text \rangle in the sense of The \textit{TeX}book pp. 275f. and 385; i.e., for two tokens $a$, $b$ and a token list $\beta$, $(a) \circ \beta \circ (b)$ is a brace group exactly if $a$ is from $X_1$, $b$ is from $X_2$, and $\beta$ is \langle balanced text \rangle. The conditions are more formal than what I can find in The \textit{TeX}book, but still they don’t give me an idea of all possibilities. This should be improved by the following recursive definition:

\begin{verbatim}
\DoWith\typein$#^{.a_1}$\StopDoing
\end{verbatim} as to what works. (The claim is not affected by one or two surprises.)\footnote{Moreover, \DoWith\typein$\{a_1}\StopDoing$ tells something about “parameter tokens.”} Characters with different category codes either are not converted into a character token or are not accepted as macro arguments. The latter applies to “brace” tokens in $X_1$, $X_2$ and to the single space token \texttt{.10}.

\footnote{If you use \DoWithAllOf\typein{a_a\{a(a)}} instead, the entire token sequence \texttt{8} will appear in the token buffer “at once.”}
B1. The empty list is balanced text. B2. For any token \( t \) not in \( X_1 \) or \( X_2 \), the single-item token list \( (t) \) is balanced text. (Such a token is either a \textit{named} token or a \textit{character} token from \( X_E \) or the space token \( \text{space} \).) B3. If \( \alpha \) and \( \beta \) are balanced texts, then \( \alpha \circ \beta \) is balanced text. B4. If \( \beta \) is balanced text, \( a \) is from \( X_1 \), and \( b \) is from \( X_2 \), then \( (a) \circ \beta \circ (b) \) is balanced text. (This is a brace group, and the only way of getting a brace group.) B5. Nothing else is balanced text.

In other words, a token list is a brace group if and only if it is balanced text and starts with a token from \( X_1 \) and ends with a token from \( X_2 \).\(^{23}\)

1.9 Summary: “Commands” Usable with dowith

In the implementation section, you learn about \( \text{\DoWith}(\text{cmd}), \text{\DoWithAllOf}(\text{cmd}), \text{\DoWithAllIn}(\text{cmd}) \).

(\TeX{} users may type \{\text{\DoWith}(\text{cmd})\} instead.) What \text{\DoWith}(\text{cmd})s are allowed?

1. All \textbf{one-parameter macros} \text{\DoWith}(\text{cmd}) work this way, unless there are programming mistakes outside dowith (also thinking of arguments that take over control from dowith commands before the argument list is finished).

2. \textbf{Other one-parameter “commands”} \text{\DoWith}(\text{cmd}) such as \TeX{} \textbf{primitives} may work—you must think of the fact that surrounding \textit{braces} are added.\(^{24}\) So the primitives \texttt{\hbox} and \texttt{\vbox} work, for instance. \texttt{\show} is an example that doesn’t work at all, it takes the single starting brace token and then confuses \DoWith.

3. Some \text{\DoWith}(\text{cmd})s taking \textbf{no argument} may make sense, e.g., for getting

(a) apples,
(b) pears,
(c) peaches

from

\begin{verbatim}
\begin{enumerate}
  \DoWithAllOf{\item}{apples, pears, peaches}
\end{enumerate}
\end{verbatim}

Recall that \texttt{\item} at most takes an \textit{optional} argument.

4. \text{\DoWith}(\text{cmd}) must \textbf{not take more than one} parameter. A different package will support multi-parameter macros.

\(^{23}\)Again, this may be more of a draft for a paper, or notes for it, than package documentation.

\(^{24}\)\textit{TODO}: in the future, variants not adding braces could be added.
2 Similar Commands in other Packages

2.1 “Heavy” Packages
The ε-TeX-related packages etools (Florent Chervet), etoolbox (Philipp Lehman), and texapi (Paul Isambert) seem to include and (very much) extend the functionality of dowith. Also the \ForEach... macros of forarray (Christian Schröppel) seem to extend the present \DoWith... commands. Moreover, Ahmed Musa describes such commands as “Parsing ‘tsv’ lists” in documenting his \catoptions package. moredefs (Matt Swift) provides list handling commands like the few that are here.\footnote{arrayjobx provides somewhat “exotic” handling of “lists.”} —In October 2012, Ahmed Musa’s loops appeared on CTAN, offering loops of several “categories” about as those that are listed below, very elaborate.—I do not want to load that much. I need and only need something excessively simple, very few lines of code, as presented in Section 3. The next sections somewhat point out single features of loop constructs that I do not want to have.

2.2 Separators
Regarding LATEX macros in \texttt{latex.ltx}, the basic macro \DoWith of the present package resembles \texttt{@for} very much, which likewise deals with lists without separators. By contrast, LATEX’s \texttt{\@for} deals with \emph{comma-separated} lists (such as lists of package options). With comma-separated lists, a “string” of characters counts as an item when it is delimited by commas, or by a comma and the list “border,” or spaces may be used as separators additionally. However, when LATEX analyzes such lists (in “TEX’s gullet”), it uses representations by \emph{character tokens} of them.

The more recent \texttt{lmake} (Shengjun Pan) provides a key-value syntax for printing lists of complex mathematical expressions easily (using some assignments) as well as defining commands according to a pattern from a list. Those lists are comma-separated.

2.3 “For” Loops vs. “Foreach” Loops
What about forloop (Nick Setzer), \texttt{multido} (Timothy Van Zandt, Rolf Nieprakhsch, Herbert Voß), and xfor (Nicola Talbot)?\footnote{arrayjobx provides somewhat “exotic” handling of “lists.”} xfor is just a reimplementation of \texttt{\@for}. forloop and multido are more close to “real ‘for’ loops” (cf. Wikipedia). Loops of the latter kind go through a certain set as well, but such sets rather consist of \emph{numbers} and are exhausted by incrementing (or also decrementing) variables (counters). This is essentially not needed (neither helpful) when a list literally is \emph{enumerated}—such loops are distinguished as “foreach loops.”
2.4 Iterators

So \texttt{DoWith} and \texttt{@tfor} rather provide “foreach” loops. A major difference between them is that the latter uses a “loop variable” or “iterator” to which the elements of the list are assigned. \texttt{DoWith}\texttt{(cmd)} does not use such a loop variable or such assignments and thus is “expandable” at least when \texttt{(cmd)} (and the elements, depending on \texttt{(cmd)}) are expandable. On the other hand, \texttt{@tfor} applies some procedure to the list elements without needing a name for the procedure (or a macro storing the procedure). I wondered whether behind \LaTeX{}’s \texttt{@tfor} (and \texttt{@for}) there was an “ideological” consideration such as “A loop must have a loop variable!” …

Hopefully more clearly on “loop variable” vs. our approach: In order to run \texttt{(code-before)\langle item\rangle (code-after)} on each \texttt{\langle item\rangle} of a \texttt{\langle list\rangle}, we here define \texttt{\do{}} as \texttt{\#1 \rightarrow \langle code-before\rangle \#1 \langle code-after\rangle}\texttt{.} (13)

and then run \texttt{\do{\langle item\rangle}} for each \texttt{\langle item\rangle} in \texttt{\langle list\rangle}\texttt{.} (14)

(\texttt{\do{}} is only an example command that \texttt{dowith} supports especially.) In \texttt{latex.ltx} instead, we find things like

\texttt{\@tfor\@tmp:=\langle list\rangle \do{\langle code-before\rangle \@tmp \langle code-after\rangle}} (15)

where \texttt{\@tmp} is a \texttt{macro} that is set to be \texttt{\langle item\rangle} at each iteration of the loop, by

\texttt{\def\@tmp{\langle item\rangle}} (16)

within \texttt{\@tfor} loop. After that,

\texttt{\langle code-before\rangle \@tmp \langle code-after\rangle}} (17)

from 15 is run.\texttt{\@tmp} like 15 is stored in a larger macro. \texttt{\do{}} in 15 does not act as a macro, it just delimits a macro parameter in order give a feeling of some familiar programming structure. This organisation of macros is fine when the loop body code is only used by the containing macro, while the \texttt{dowith} approach to store the “loop body” in an own macro has been useful when the loop body code also is used for different purposes or when it has been introduced before I thought of using it in a loop.

Note that this only was an example. In general, \texttt{\langle item\rangle} may appear more than once in the “loop body.”

“Expandability” by avoiding something iterating \texttt{\def\@tmp{\langle item\rangle}} and doing iteration in \LaTeX{}’s gullet (\texttt{\do{}} or so must have been defined earlier) is essential especially within \texttt{\write{}}. Assignments do not work there. A major motivation for developing \texttt{dowith} developed with the \texttt{blog} package that \texttt{\writes{}} HTML code. Assignments happen in “\LaTeX{}’s stomach.” That place might be called the “instruction buffer” to which the “expansion processor” moves items from the incoming token buffer that cannot be expanded (any more).

\footnote{Cf. description of procedure in terms of tokens in Section 1.7}
2.5 Separator Macros

Commands like \DoWith also could save tokens thinking of list macros (in \LaTeX{}\texttt{latex.ltx}) that use a separator macro which may be used as a command to be applied to the list elements. One example is \dospecials that already is in Plain \TeX{} and expands to

\begin{verbatim}
  \do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do
\end{verbatim}

An important application of \dospecials is temporarily switching off the “special” functionality of the “elements” in \dospecials. With \LaTeX{}, this may happen thus:

\begin{verbatim}
  \let\do@makeother\dospecials
\end{verbatim}

With dowith, you can do the same with a shorter variant \specials of \dospecials, defined by

\begin{verbatim}
  \def\specials{\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do\do}
\end{verbatim}

and then

\begin{verbatim}
  \DoWithAllIn@makeother\specials
\end{verbatim}

\texttt{latex.ltx} uses \@elt instead of \do for its own list macros.

2.6 Ye Olde \textbackslash loop

There also is \texttt{\[\texttt{\textbackslash loop}(\texttt{\textbackslash loop-body})\texttt{\textbackslash repeat}\]} in Plain \TeX{} and a refined\textsuperscript{27} version of it in \texttt{latex.ltx}. It is not expandable since it starts with an assignment for \texttt{\textbackslash body} (Plain \TeX{}) or \texttt{\textbackslash iterate} (\texttt{latex.ltx}), and then some assignments are needed to stop the loop, such as incrementing or decrementing a counter. As to the programming structure, it is very simple and general, I think any kind of loop can be implemented by this (apart from nested loops). E.g., I realize\textsuperscript{28} that even a “foreach” loop could be implemented by managing a list macro, e.g., using \LaTeX{}’s internal \@next.

2.7 Without Iterator and Separators

In \LaTeX{}’s \texttt{tools} bundle, \xspace was developed in the nineties by David Carlisle. It had a rather fixed exception list implemented by a deeply nested conditional. In 2004 Morton Høgholm joined, and now \xspace has a list macro \texttt{\@xspace@exceptions@tlp} without separators. It is handled like here, except that it “breaks” the loop when an item is found that applies. After the “next” token is stored by the usual \texttt{\futurelet}, the exception list is searched without using an iterator. Addition and removal commands are provided as well.

\textsuperscript{27}Using Kabelschacht’s suggestion, cf. Section 4

\textsuperscript{28}2012-05-20
3 Implementation

3.1 Package File Header (Legalese)

\def\filename{dowith} \def\fileinfo{simple list loop (UL)}
\def\filedate{2012/11/05} \def\fileversion{v0.3}

%% Copyright (C) 2011 2012 Uwe Lueck,
%% http://www.contact-ednotes.sty.de.vu
%% -- author-maintained in the sense of LPPL below --
%%
%% This file can be redistributed and/or modified under
%% the terms of the LaTeX Project Public License; either
%% version 1.3c of the License, or any later version.
%% The latest version of this license is in
%% http://www.latex-project.org/lppl.txt
%% We did our best to help you, but there is NO WARRANTY.
%%
%% Please report bugs, problems, and suggestions via
%%
%% http://www.contact-ednotes.sty.de.vu

3.2 Proceeding without \LaTeX

v0.3 mainly replaces imitating the german.sty approach to genericity by plainpkg.

\input plainpkg
\ProvidesPackage{dowith}[2012/11/05 simple list loop (UL)]
\PushCatMakeLetterAt

If \LaTeX{} is not present ...

\ifltx \else

... an old version of its \texttt{\textbackslash in@} is introduced. It is bad as a subword test (false positive cases, cf. \texttt{\textbackslash fifinddo} documentation), but \texttt{\textbackslash dowith} will check for single tokens only. If \LaTeX{} is present, on the other hand, \texttt{\textbackslash in@} is recognized while skipping \texttt{false} parts of conditionals, without being matched by some \texttt{\textbackslash fi} before the next \texttt{\textbackslash else}, so I hide it by \texttt{\textbackslash csname}:

\expandafter\newif\csname ifin@\endcsname
\def\in@{\if\in@false\else\in@true\fi}
\if\in@\else

\fi
3 IMPLEMENTATION

3.3 Applying a Command

3.3.1 Core

\DoWith{\langle cmd \rangle}{\langle list \rangle}\StopDoing applies \langle cmd \rangle to all elements of \langle list \rangle. An element of \langle list \rangle (after tokenizing) may be either a single token or a group \langle balanced \rangle.

\def\DoWith#1#2{% 
\ifx\StopDoing#2\empty 
  The previous \empty (replacing \%) is a bug fix as of v0.22 (June 2012), while in my extension draft I already had it in January 2012. It allows “empty” arglist items ‘\{1\}2’. Before v0.22, such an empty brace group would have resulted in comparing \StopDoing with \else, so ‘\{1\}2’ would have had the same effect as \StopDoing, the token text after \else until \fi would have been skipped. Instead, the user may have a reason to allow empty arguments/brace groups.

\else#1{#2}\expandafter\DoWith\expandafter#1\fi}

\StopDoing delimits the list:

\let\StopDoing\DoWith

\let\StopDoing*

\DoWithAllOf{\langle cmd \rangle}{\langle list \rangle}\StopDoing:

\def\DoWithAllOf#1#2{\DoWith#1#2\StopDoing}

3.3.2 \do being the Command

When the \langle list \rangle is worked at a single time in the \TeX run where assignments are possible, instead of introducing a new macro name for \langle cmd \rangle you can use \do for \langle cmd \rangle as a “temporary” macro and define it right before

\DoWith{\do}{\langle list \rangle}\StopDoing

However, we provide

\DoDoWith{\langle cmd \rangle}{\langle list \rangle}\StopDoing

as a substitute for the former line that at least saves one token. For the definition of \do, we provide \setdo{\langle def-text \rangle}\ DoDoWith{\langle cmd \rangle}{\langle list \rangle}\StopDoing.

\renewcommand{\do}[1]{\langle def-text \rangle},

so \langle def-text \rangle should contain a #1:
With \letdo{cmd} that is provided next where \textit{cmd} is defined elsewhere, you could type
\letdo{cmd}\DoDoWith\StopDoing

It seems to me, however, that you better type
\dowith{cmd}{list}\StopDoing

instead. So I provide \letdo although I consider it useless here. It is provided somewhat for the sake of “completeness,” thinking that it might be useful at other occasions such as preceding \dospecials.

\def\letdo{\let\do}

\DoDoWith has been described above:

\def\DoDoWithAllOf{\DoWithAllOf\do}

By analogy to \DoWithAllOf, we provide \DoDoWithAllOf{\langle list \rangle}:

\def\DoDoWithAllIn#1#2{\expandafter\DoWith\expandafter#1#2\StopDoing}

\DoDoWithAllIn{\langle list-macro \rangle}{\langle cmd \rangle} saves a backslash or token for \do as above in Sec. 3.3.2.

3.4 Handling List Macros

3.4.1 Initializing

Here is some advanced \let{cmd}\empty, perhaps a little irrelevant for practical purposes. Both
\InitializeListMacro{\langle list-macro \rangle}
3 IMPLEMENTATION

and

\ReInitializeListMacro{\langle list-macro\rangle}

attempt to “empty” \langle list-macro\rangle, and when we don’t believe that \LaTeX{} has been loaded, both do the same indeed. Otherwise the first one complains when \langle list-macro\rangle seems to have been used earlier while the second complains when \langle list-macro\rangle seems not to have been used before:

\begin{verbatim}
42 \ifltx \%\ v0.3
43 \def\InitializeListMacro#1{\@ifdefinable#1{\let#1\empty}}
44 \def\ReInitializeListMacro#1{\
45 \edef\@tempa{\expandafter\@gobble\string#1}\
46 \expandafter\@ifundefined\expandafter{\@tempa}\
47 \{\@latex@error{\noexpand#1undefined}\@ehc}\%
48 \{\let#1\empty\}}
49 \else \def\InitializeListMacro#1{\let#1\empty} \% not \@empty 2011/11/07
50 \let\ReInitializeListMacro\InitializeListMacro
51 \fi
\end{verbatim}

\ToListMacroAdd{\langle list-macro\rangle}{\langle cmd-or\rangle} appends \langle cmd-or\rangle to the replacement token list of \langle list-macro\rangle. \langle cmd-or\rangle may either be tokenized into a single token, or it is some \{\langle balanced\rangle\}.

\begin{verbatim}
53 \def\ToListMacroAdd#1#2{\DefExpandStart#1{#1#2}}
54 \def\DefExpandStart#1{\expandafter\def\expandafter#1\expandafter}
\end{verbatim}

3.4.2 Testing for Occurrence of a Token

\TestListMacroForToken{\langle list-macro\rangle}{\langle cmd\rangle} sets \@in@true when \langle cmd\rangle occurs in \langle list-macro\rangle and sets \@in@false otherwise:

\begin{verbatim}
55 \def\TestListMacroForToken#1#2{\%
56 \expandafter \iny@ \expandafter \#2\expandafter\expandafter#1\expandafter}\%
\end{verbatim}

Indeed I removed an earlier \IfTokenInListMacro, now it’s a kind of compromise between having a shorthand macro below and a generalization for users of the package.

3.4.3 Adding and Removing

\FromTokenListMacroRemove{\langle list-macro\rangle}{\langle cmd\rangle} removes the token corresponding to \langle cmd\rangle from the list stored in \langle list-macro\rangle (our parsing method does not work with braces):

\begin{verbatim}
57 \def\FromTokenListMacroRemove#1#2{\%
\end{verbatim}

I am not happy about defining two parser macros, but for now...
3 IMPLEMENTATION

\TestListMacroForToken#1#2\%\ifin@\Def\RemoveThisToken##1#2\%\expandafter\DefExpandStart\expandafter#1\expandafter{\%\expandafter\RemoveThisToken#1}\%\fi}

TODO warning otherwise?

... but this only removes a single occurrence ...

\InTokenListMacroProvide{(list-macro)}{(cmd)} avoids multiple entries of a token by not adding anything when (cmd) already occurs in (list-macro) (again, this does not work with braces, try \in@{}}{{}{{}}).

\def\InTokenListMacroProvide#1#2{\TestListMacroForToken#1#2\%\ifin@\else\fi\ToListMacroAdd#1#2\%\fi}

3.5 Leaving and History

\PopLetterCatAt \endinput

VERSION HISTORY

v0.1 2011/06/23/28 stored separately
v0.2 2011/11/02 simpler, documented
v0.21 2012/05/14 fix for "generic" and 'typeoutfileinfo': @ before ...!

2011/11/19 modified LaTeX supplements
v0.21a 2012/05/19 \labels sec:apply, sec:core; \pagebreak?
v0.22 2012/06/04 allow {} items
v0.3 2012/11/05 updating copyright, using 'plainpkg', rewording documentation there
4 Ack.: 25 Years of Kabelschacht’s \expandafter

The essential idea of \texttt{dowith} and \texttt{DoWith} is

\begin{verbatim}
\if\texttt{code}\expandafter\expandafter\expandafter\expandafter\expandafter\expandafter\texttt{expandafter}\expandafter\texttt{one-token}\texttt{fi}
\end{verbatim}

It was described by Alois Kabelschacht as \texttt{\expandafter vs. \let and \def in Conditionals and a Generalization of \texttt{PLAIN’s \texttt{\loop}}} in \texttt{TUG}boat Vol. 8 (1987), No. 2, pp. 184f. (a little more than one column).\footnote{http://tug.org/TUGboat/tb08-2/tb18kabel.pdf} See some German biographical notes on Kabelschacht in the German Wikipedia.\footnote{http://de.wikipedia.org/wiki/Benutzer:RolteVolte/Alois_Kabelschacht} It seems to me that Knuth didn’t note this application of \texttt{\expandafter} in \texttt{The \TeX{}book}.\footnote{However, the paper ‘uses the fact that the expansion of both \texttt{\else \texttt{\fi and \texttt{\fi is empty.’ In \texttt{The \TeX{}book} I only find ‘The “expansion” of a conditional is empty’ on page 213.} It was then applied in many macros of \texttt{latex.ltx}, cf. source2e.pdf.