dowith.sty

—

Apply Command to Each Item in a List of Arguments in "TEX's Mouth"*

Uwe Lück†

June 4, 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. Applications in mind belonged to LATEX, but the package should work with other formats as well. Loop and list macros in other packages are discussed. Iteration is implemented within "TEX's mouth," so works within `\write` as with blog.sty. There is no need for ε-TEX.

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

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

Contents

*This document describes version v0.22 of dowith.sty as of 2012/06/04.

†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 data base).[1]

With LaTeX, you load dowith.sty (as usually) by

```
\usepackage{dowith}
```

below the `\documentclass` line(s) and above `\begin{document}`. However, the package can also be used with other formats, just

```
\input dowith.sty
```

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.

[1] `http://www.tex.ac.uk/cgi-bin/texfaq2html?label=inst-wlcf`

First of all, we are not referring to LaTeX `list` environments such as `enumerate` or `itemize`; neither to "TODO" lists of what needs to be done soon.

Rather, dowith allows you to abbreviate

⟨*cmd*⟩⟨*arg-1*⟩⟨*cmd*⟩⟨*arg-2*⟩ . . . ⟨*cmd*⟩⟨*arg-n*⟩

by

`\DoWith`⟨*cmd*⟩⟨*arg-1*⟩⟨*arg-2*⟩ . . . ⟨*arg-n*⟩`\StopDoing`

or by

`\DoWithAllOf`⟨*cmd*⟩`{`⟨*arg-1*⟩⟨*arg-2*⟩ . . . ⟨*arg-n*⟩`}`

With small n, one may doubt whether this really is an abbreviation . . . ; anyway,

⟨*arg-1*⟩⟨*arg-2*⟩ . . . ⟨*arg-n*⟩

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

⟨*arg-1*⟩, ⟨*art-2*⟩, . . . , ⟨*arg-n*⟩

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

We might say that `aa` is a list of *two* items, ⟨*arg-1*⟩ being `a` and ⟨*arg-2*⟩ being `a`, too.

When we do *three* keystrokes to get `a␣a` instead of `aa`, we still have *two* items, ⟨*arg-1*⟩ being `a` and ⟨*arg-2*⟩ being `a` too. Strange, isn't it?

Also, when in `aa` we replace the first `a` by a backslash, `\`, we get `\a`, and this is a list of a *single* item, ⟨*arg-1*⟩ = `\a` . . .

You shouldn't believe these stories of mine entirely. What I am alluding to is that the *"items"* dowith is about are determined in terms of *TeX's tokens*, and the relation between the "characters you type" and *TeX's tokens* is not entirely straightforward.

1.3 The Notion of Arglists for LaTeX Users

Still, it may suffice to clarify what counts as an ⟨*arg-i*⟩ without speaking of *tokens* explicitly: It is simply what a *one-parameter macro* (where the parameter is *not delimited* in terms of *The TeXbook* pp. 203f.) can take as an *argument.*

The lists dowith is about then are lists *of possible arguments* in the previous sense—let me call them *"arglists."*[2] The single *items* of such lists are those single possible arguments. They become *actual* arguments beginning from the leftmost *possible* one when dowith presents them to that ⟨*cmd*⟩ mentioned earlier—where ⟨*cmd*⟩ *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 ⟨*arg-i*⟩ just has form `{`⟨*ark-i*⟩`}`, i.e., ⟨*arg-i*⟩ = `{`⟨*ark-i*⟩`}` for some

[2] Not to be confused with German Arglist.

⟨*ark-i*⟩. Such arguments are also called *"brace groups"*. (LaTeX's *optional* arguments `[`⟨*extra*⟩`]` do not count as possible arguments here, they are not macro arguments in the sense of *The TeXbook*.) In this restricted LaTeX sense, arglists consist of brace groups

`{`⟨*ark-1*⟩`}{`⟨*ark-2*⟩`}` ... `{`⟨*ark-n*⟩`}`,

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 TeX's Tokens

What TeXnically matters is what happens in "TeX's mouth,"[3] as some authors have suggested a metaphor,[4] or somewhere deeper. The dowith package is a tool to control those events (and actually, it is confined to TeX's mouth).

The "characters you type" are somewhere in front of "TeX's mouth", while *in* TeX's mouth, there are *tokens*. Before TeX *swallows* them, it often manipulates them in some ways, after they got *into* its mouth.

More formally, TeX has a *character buffer*. It forms a single token from an initial segment of the buffer content—unless there is a special situation with blank spaces or something pathological. When an *escape character*, as the backslash usually is one, has been noticed recently (that isn't followed by another one immediately), the character buffer may need to be feeded from more outside, until it contains enough material to form a token from. The character(s) *after* the escape character until some delimiting character form a *string* that is the *name* of the token that is formed. What has been used to form a token is removed from the character buffer.

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 mouth (perhaps unless one considers a one-step macro expansion a two-or-more-step procedure).

For every *string of characters*, there is exactly one (possible) *named token* whose name the string is.[5] It is so common (starting from *The TeXbook*) to denote the token whose name is ⟨*string*⟩ by '\⟨*string*⟩'. For instance, the token whose name is `input` is denoted by '`\input`'. On the other hand, on page 7 of *The TeXbook* '`\input`' is a "string of characters." With this notation, it is already difficult to explain what the LaTeX command `\DeclareRobustCommand` does or what the difference between a starred LaTeX command and a starred LaTeX environment is.[6] *The TeXbook* makes it worse by saying on page 39: *"A control sequence is*

[3] Cf. *The TeXbook* p. 46.

[4] Alan Jeffrey: "Lists in TeX's Mouth," TUGboat Vol. 11 (1990), No. 2, pp. 237–245), `tug.org/TUGboat/tb11-2/tb28jeffrey.pdf`.

[5] "Possible" refers to the fact that TeX does not store named tokens anywhere before they appear in its mouth, maybe apart from "primitive" tokens that have a "pre-assigned meaning" when a TeXrun 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.

[6] A reader knowing LaTeX only thinks that '`\\␣`' is the result of typing a double backslash and a space and that '`\equation*`' is the "command" `\equation` followed by a `*`.

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 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 is 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."*

The TEXbook does offer an alternative notation for named tokens: "boxing;" so the token whose name is `input` can be denoted by the rather "graphical" notation '$\boxed{\texttt{input}}$' (used only exceptionally).[7] *I* would suggest something like '`ntok(input)`' for clarity and '`?input`' for brevity.[8]

Character tokens get into TEX's mouth by tokenization when characters begin the buffer content while *not* scanning a name for a named token. A single character then is removed from the character buffer, and a token storing its character code and current category code is pushed into TEX's mouth.

Named tokens may get into TEX's mouth by "tokenization" as described above, i.e., they are drawn from the character buffer. But they also can appear in TEX's mouth "from within," by the manipulation inside TEX's mouth.

More formally, those manipulations are called "expansion," and TEX's mouth 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 mouth that emerged from tokenization a "long time ago", maybe in a previous run that created the *format* (TEX's variant `INITEX`); or tokens may appear by some hardwired expansion function.

However, *named tokens* may get into TEX's mouth also by *expansion*, never having been drawn by tokenization and not being hardwired. This happens by the `\csname` name construct. The input *code* may contain

```
\csname␣tupni\endcsname
```

This may be converted into 7 tokens entering TEX's mouth, the first one being `ntok(csname)`, the last one `ntok(endcsname)`, and five character tokens in between. Due to some *function* (which I would denote as `*csname`) originally associated with the token `ntok(csname)`, those seven tokens then are replaced by `ntok(tupni)`, the named token whose name is `tupni`. It is not required that the TEX program knows about a token `ntok(tupni)`, neither anybody must type '`\tupni`' in any file.[9]

1.5 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

```
a␣a\a{a}                                                    (1)
```

[7] The box notation is introduced on page 38 without explanation, as if it explained something.

[8] I am suggesting the question mark for named tokens since TEX "must look up the current definition" of a named token according to *The TEXbook* p. 39, while the meaning of character tokens rather is "fixed," at least according to *The TEXbook* p. 39. However, *active*-character such as .~ are in the same situation as named tokens as to this respect. The dot notation may be fine for them, though.

[9] These considerations may not be essential here, rather a draft for a paper. Using `dowith`, one better just thinks of the arglist items one actually lists.

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

$$\mathtt{a}_{11}\ \mathtt{␣}_{10}\ \mathtt{a}_{11}\ \mathtt{ntok(a)}\ \mathtt{\{}_1\ \mathtt{a}_{11}\ \mathtt{\}}_2 \tag{2}$$

—with notation from Section 1.4 and *The TEXbook*'s notation $\langle char\rangle_{\langle cat\rangle}$ for the *character token* that TEX's tokenizer forms from $\langle char\rangle$ in the character buffer when $\langle char\rangle$'s category code is $\langle cat\rangle$.

It turns out that the token list in 2 provides an arglist of *four* items: The token $\mathtt{a}_{11}$ at the first and third place, the named token `ntok(a)`, and the entire token list $\mathtt{\{}_1\mathtt{a}_{11}\mathtt{\}}_2$ as a single item—a "brace group." The space token is ignored.[10]

You can try this after `\renewcommand{\a}{A}`[11] with dowith:

$$\mathtt{\backslash DoWith\backslash typein␣a␣a\backslash a\{a\}\backslash StopDoing} \tag{3}$$

Then LATEX shows `a`, `a`, `A` from `\a`, and another `a` from within the braces—`\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 "commma-separated lists" within brackets may clarify the difference (that the juxtaposition notation tends to conceal). The *token* list is

$$(\mathtt{a}_{11}, \mathtt{␣}_{10}, \mathtt{a}_{11}, \mathtt{ntok(a)}, \mathtt{\{}_1, \mathtt{a}_{11}, \mathtt{\}}_2) \tag{4}$$

while the list of macro arguments is

$$(\mathtt{a}_{11}, \mathtt{a}_{11}, \mathtt{ntok(a)}, (\mathtt{\{}_1, \mathtt{a}_{11}, \mathtt{\}}_2)). \tag{5}$$

2 or 4 simply is *not* an arglist (since neither $\mathtt{\{}_1$ nor $\mathtt{\}}_2$ 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(*!?*).

1.6 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 $\langle char\rangle$, i.e., adding the *standard* category code as in *The TEXbook* (page 37). Then 2 would read[12]

$$\mathtt{.a\,.␣\,.a\,?a\,.\{\,.a\,.\}} \tag{6}$$

and the corresponding arglist is

$$(\mathtt{.a}, \mathtt{.a}, \mathtt{?a}, (\mathtt{.\{\,.a\,.\}})) \tag{7}$$

In "retrospect," the result of tokenizing 2 should be

$$\mathtt{?DoWith\,?typein\,.a\,.␣\,.a\,?a\,.\{\,.a\,.\}\,?StopDoing} \tag{8}$$

[10] *The TEXbook* p. 201: "TEX doesn't use single spaces as undelimited arguments."
[11] Otherwise `\a` is a one-parameter macro that breaks dowith's control.
[12] See Section 1.4 for the question mark.

and the intention is that it works like

?typein.{.a.} ?typein.{.a.} ?typein.{?a.} ?typein.{.a.} (9)

(The definition of \DoWith in Section 3.3.1 indeed adds surrounding braces, if missing.) However, TEX rather tries to work with as few tokens ahead as possible. When it finds ?DoWith and the latter's meaning is the one intended by dowith, it first looks for nothing more than the two arguments required by our definition of \DoWith. A few moments later, the token buffer's content will just be[13]

?typein.{.a.} ?expandafter ?DoWith ?expandafter ?typein ?fi (10)

Next ?typein.{.a.} is expanded according to the code for \typein in latex.ltx. Some unexpandable tokens will emerge and be moved into the "command 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

?DoWith ?typein (11)

Another token is ordered from the tokenizer to provide a second argument for expanding ?DoWith. The token .␣ comes 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 ...

1.7 Summary of Possible Arglist Items

For $0 \le i \le 15$, 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 $\}_2$.

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 $\mathrm{X}_E = \bigcup_{i \in E} \mathrm{X}_i$).

Let $\circ$ be the *concatenation* operation among token lists.[14]

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 λ than X_1 occurrences in ρ ("don't close before opening").

The second claim can be checked with

\DoWith\typein$#^_a1~\StopDoing (12)

as to what works. (The claim is not affected by one or two surprises.)[15] Characters with different category codes either are not converted into a character token[16] or are not accepted

[13] If you use \DoWithAllOf\typein{a␣a\a{a}} instead, the entire token sequence 8 will appear in the token buffer "at once."

[14] 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.

[15] Moreover, \DoWith\typein#1\StopDoing tells something about "parameter tokens."

[16] *The TEXbook* p. 47.

as macro arguments. The latter applies to "brace" tokens in X_1, X_2 and to the single space token $\sqcup 10$.

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

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 *named* token or a *character* token from X_E or *the space token* $\sqcup 10$.) B3. If α and β are balanced texts, then $\alpha \circ \beta$ is balanced text. B4. If β 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.[17]

1.8 Summary: "Commands" Usable with dowith

In the implementation section, you learn about

`\DoWith`⟨*cmd*⟩, `\DoWithAllOf`⟨*cmd*⟩, and `\DoWithAllIn`⟨*cmd*⟩.

(LaTeX users may type `{`⟨*cmd*⟩`}` instead.) What ⟨*cmd*⟩s are allowed?

1. All **one-parameter macros** ⟨*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. **Other one-parameter** "commands" ⟨*cmd*⟩ such as TeX **primitives** may work—you must think of the fact that surrounding *braces* are added.[18] So the **primitives** `\hbox` and `\vbox` work, for instance. `\show` is an example that doesn't work at all, it takes the single starting brace token and then confuses `\DoWith`.
3. Some ⟨*cmd*⟩s taking **no argument** may make sense, e.g., for getting
   (a) apples,
   (b) pears,
   (c) peaches

   from

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

   Recall that `\item` at most takes an *optional* argument.
4. ⟨*cmd*⟩ must **not take more than one** parameter. A different package will support multi-parameter macros.

[17] Again, this may be more of a draft for a paper, or notes for it, than package documentation.
[18] 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 etextools (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.[19] (I do not want to load that much.)

2.2 Separators

Regarding LaTeX macros in `latex.ltx`, the basic macro `\DoWith` of the present package resembles `\@tfor` very much, which likewise deals with lists without separators. By contrast, LaTeX's `\@for` deals with *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 mouth"), it uses representations by *character tokens* of them.

The more recent 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), multido (Timothy Van Zandt, Rolf Niepraksch, Herbert Voß), and xfor (Nicola Talbot)?

xfor is just a reimplementation of `\@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 *numbers* and are exhausted by incrementing (or also decrementing) variables (counters). This is essentially not needed (neither helpful) when a list literally is *enumerated*—such loops are distinguished as "foreach loops."

2.4 Iterators

So `\DoWith` and `\@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. `\DoWith`⟨*cmd*⟩ does not use such a loop variable or such assignments and thus is "expandable" at least when ⟨*cmd*⟩ (and the elements, depending on ⟨*cmd*⟩) are expandable. On the other hand,

[19] arrayjobx provides somewhat "exotic" handling of "lists".

`\@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 `\@tfor` (and `\@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

⟨*code-before*⟩⟨*item*⟩⟨*code-after*⟩

on each ⟨*item*⟩ of a ⟨*list*⟩, *we here*

define `\do` as `#1` → ⟨*code-before*⟩`#1`⟨*code-after*⟩ (13)

and then run `\do{`⟨*item*⟩`}` for each ⟨*item*⟩ in ⟨*list*⟩,[20]

always replacing `\do{`⟨*item*⟩`}` by `\do{`⟨*item*⟩`}\do`. (14)

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

`\@tfor\@tmp:=`⟨*list*⟩`\do{`⟨*code-before*⟩`\@tmp`⟨*code-after*⟩`}` (15)

where `\@tmp` is a *macro* that is set to be ⟨*item*⟩ at each iteration of the loop, by

`\def\@tmp{`⟨*item*⟩`}` (16)

within `\@tforloop`. After that,

⟨*code-before*⟩`\@tmp`⟨*code-after*⟩ (17)

from 15 is run.—17 like 15 is stored in a larger macro. `\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 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, ⟨*item*⟩ may appear more than once in the "loop body."

"Expandability" by *avoiding* something iterating `\def\@tmp{`⟨*item*⟩`}` and doing iteration in TeX's mouth (`\do` or so must have been defined earlier) is essential especially within `\write`. Assignments do not work there. A major motivation for developing dowith developed with the blog package that `\write`s HTML code. Assignments happen somewhere *behind* "TeX's mouth." That place might be called the "command buffer" to which the "expansion processor" moves items from the incoming token buffer that cannot be expanded (any more).

[20] Cf. description of procedure in terms of tokens in Section 1.6.

2.5 Separator Macros

Commands like `\DoWith` also could save tokens thinking of list macros (in LaTeX/`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

```
\do\␣\do\\\do\{\do\}\do\$\do\&\do\#\do\^\do\_\do\%\do\~
```

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

```
\let\do\@makeother\dospecials
```

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

```
\def\specials{\␣\\\{\}\$\&\#\^\_\%\~}
```

and then

```
\DoWithAllIn\@makeother\specials
```

`latex.ltx` uses `\@elt` instead of `\do` for its own list macros.

2.6 Ye Olde `\loop`

There also is `\loop`⟨*loop-body*⟩`\repeat` in Plain TeX and a refined[21] version of it in `latex.ltx`. It is *not* expandable since it starts with an assignment for `\body` (Plain TeX) or `\iterate` (`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[22] 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 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 `\@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 `\futurelet`, the exception list is searched without using an iterator. Addition and removal commands are provided as well.

[21] Using Kabelschacht's suggestion, cf. Section 4
[22] 2012-05-20

3 Implementation

3.1 Package File Header (Legalese)

```
\def\filename{dowith}     \def\fileinfo{simple list loop (UL)}
\def\filedate{2012/06/04} \def\fileversion{v0.22}
%% Copyright (C) 2011 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

A little LaTeX as in Bernd Raichle's ngerman.sty:

```
\chardef\atcode=\catcode`\@
\catcode`\@=11 % \makeatletter
\begingroup\expandafter\expandafter\expandafter\endgroup
```

I need `\ProvidesPackage` for fileinfo, my package version tools.

```
\expandafter\ifx\csname ProvidesPackage\endcsname\relax
```

When `\ProvidesPackage` is not defined, we provide a version of LaTeX's `\in@` (an old version that may wrongly claim to have found an occurrence of a sequence, but is correct for single tokens) for checking token list macros. LaTeX must not see `\ifin@` when it parses the `\ifx` conditional:

```
  \expandafter\newif\csname ifin@\endcsname
  \def\in@#1#2{%
    \def\in@@##1#1##2##3\in@@{%
      \ifx\in@##2\in@false\else\in@true\fi}%
    \in@@#2#1\in@\in@@}
```

readprov stops reading the file at `\Provides...`, therefore ...

```
  \long\def\@gobble#1{} \expandafter\@gobble
\else
  \expandafter\@firstofone
\fi
{ \ProvidesPackage{\filename}[\filedate\space
                    \fileversion\space \fileinfo] }
```

3.3 Applying a Command

3.3.1 Core

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

```
\def\DoWith#1#2{%
    \ifx\StopDoing#2\empty
```

The previous `\empty` (replacing `%`) is a bug fix as of v0.22, 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
```

... something arbitrary that is not expected to occur in a list. With

```
\let\StopDoing*
```

instead, the star would end lists.

`\DoWithAllOf{`⟨*cmd*⟩`}{`⟨*list*⟩`}` works like

`\DoWith{`⟨*cmd*⟩`}`⟨*list*⟩`\StopDoing` :

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

3.3.2 `\do` being the Command

When the ⟨*list*⟩ is worked at a single time in the TEX run where assignments are possible, instead of introducing a new macro name for ⟨*cmd*⟩ you can use `\do` for ⟨*cmd*⟩ as a "temporary" macro and define it right before

`\DoWith{\do}`⟨*list*⟩`\StopDoing`

However, we provide

`\DoDoWith{`⟨*cmd*⟩`}`⟨*list*⟩`\StopDoing`

as a substitute for the former line that at least saves one token. For the definition of `\do`, we provide `\setdo{`⟨*def-text*⟩`}`. It works similarly to

`\renewcommand{\do}[1]{`⟨*def-text*⟩`}`,

so ⟨*def-text*⟩ should contain a `#1`:

```
\def\setdo{\long\def\do##1}
```

With `\letdo`⟨*cmd*⟩ that is provided next where ⟨*cmd*⟩ is defined elsewhere, you could type

`\letdo`⟨*cmd*⟩`\DoDoWith`⟨*list*⟩`\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\DoDoWith{\DoWith\do}
```

By analogy to `\DoWithAllOf`, we provide `\DoDoWithAllOf{`⟨*list*⟩`}`:

```
\def\DoDoWithAllOf{\DoWithAllOf\do}
```

3.3.3 Expand List Macro

The former facilities may be quite useless as such a ⟨*list*⟩ will not be typed at a single place in the source code, rather the items to run ⟨*cmd*⟩ on may be collected occasionally when some routines run. The elements may be collected in a macro ⟨*list-macro*⟩ expanding to ⟨*list*⟩. So we provide

`\DoWithAllIn{`⟨*cmd*⟩`}{`⟨*list-macro*⟩`}`

(or `\DoWithAllIn`⟨*cmd*⟩⟨*list-macro*⟩). There is no need to type `\StopDoing` here:

```
\def\DoWithAllIn#1#2{%
    \expandafter\DoWith\expandafter#1#2\StopDoing}
```

`\DoDoWithAllIn{`⟨*list-macro*⟩`}` saves a backslash or token for `\do` as above in Sec. 3.3.2:

```
\def\DoDoWithAllIn{\DoWithAllIn\do}
```

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{`⟨*list-macro*⟩`}`

and

`\ReInitializeListMacro{`⟨*list-macro*⟩`}`

attempt to "empty" ⟨*list-macro*⟩, and when we don't believe that LaTeX has been loaded, both do the same indeed. Otherwise the first one complains when ⟨*list-macro*⟩ seems to have been used earlier while the second complains when ⟨*list-macro*⟩ seems *not* to have been used before:

```
\expandafter\ifx\csname @latex@error\endcsname\relax
  \def\InitializeListMacro#1{\let#1\empty} %% not \@empty 2011/11/07
  \let\ReInitializeListMacro\InitializeListMacro
\else
  \def\InitializeListMacro#1{\@ifdefinable#1{\let#1\empty}}
  \def\ReInitializeListMacro#1{%
    \edef\@tempa{\expandafter\@gobble\string#1}%
    \expandafter\@ifundefined\expandafter{\@tempa}%
        {\@latex@error{\noexpand#1undefined}\@ehc}%
        {\let#1\empty}}
\fi
```

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

```
\def\ToListMacroAdd#1#2{\DefExpandStart#1{#1#2}}
\def\DefExpandStart#1{\expandafter\def\expandafter#1\expandafter}
```

3.4.2 Testing for Occurrence of a Token

`\TestListMacroForToken{`⟨*list-macro*⟩`}{`⟨*cmd*⟩`}` sets `\in@true` when ⟨*cmd*⟩ occurs in ⟨*list-macro*⟩ and sets `\in@false` otherwise:

```
\def\TestListMacroForToken#1#2{%
    \expandafter \in@ \expandafter #2\expandafter{#1}}
```

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{`⟨*list-macro*⟩`}{`⟨*cmd*⟩`}` removes the token corresponding to ⟨*cmd*⟩ from the list stored in ⟨*list-macro*⟩ (our parsing method does not work with braces):

```
\def\FromTokenListMacroRemove#1#2{%
```

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

```
60      \TestListMacroForToken#1#2%
61      \ifin@
62        \def\RemoveThisToken##1#2{##1}%
63        \expandafter \DefExpandStart
64          \expandafter #1\expandafter {%
65            \expandafter\RemoveThisToken #1}%
```

TODO warning otherwise?

```
66      \fi}
```

. . . 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@{{}}{{}}`).

```
67  \def\InTokenListMacroProvide#1#2{%
68      \TestListMacroForToken#1#2%
69      \ifin@ \else                    %% TODO warning?
70          \ToListMacroAdd#1#2%
71      \fi}
```

3.5 Leaving and History

```
72  \catcode`\@=\atcode
73  \endinput
74
75  VERSION HISTORY
76  v0.1    2011/06/23/28  stored separately
77  v0.2    2011/11/02     simpler, documented
78          2011/11/03     corrected \if/\else for init
79          2011/11/07     \TestListMacroForToken, \InListMacroProvide;
80                         doc.: \pagebreak, structure
81          2011/11/19     modified LaTeX supplements
82  v0.21   2012/05/14     fix for "generic" and `typeoutfileinfo':
83                         @ before ...!
84  v0.21a  2012/05/19     \labels sec:apply, sec:core; \pagebreak?
85  v0.22   2012/06/04     allow {} items
86
```

4 Ack.: 25 Years of Kabelschacht's \expandafter

The essential idea of dowith and \DoWith is

\if⟨*code*⟩\expandafter⟨*one-token*⟩\fi

It was described by Alois Kabelschacht as "\expandafter vs. \let and \def in Conditionals and a Generalization of PLAIN's \loop" in TUGboat Vol. 8 (1987), No. 2, pp. 184f. (a little more than one column).[23] See some German biographical notes on Kabelschacht in the German Wikipedia.[24] It seems to me that Knuth didn't note this application of \expandafter in *The TEXbook*.[25] It was then applied in many macros of latex.ltx, cf. source2e.pdf.

[23] tug.org/TUGboat/tb08-2/tb18kabel.pdf

[24] de.wikipedia.org/wiki/Benutzer:RolteVolte/Alois_Kabelschacht

[25] However, the paper *'uses the fact that the expansion of both* \else ... \fi *and* \fi *is empty.'* In *The TEXbook* I only find *'The "expansion" of a conditional is empty'* on page 213.