The \LaTeX3 Interfaces

The \LaTeX3 Project*

Released 2020-03-06

Abstract

This is the reference documentation for the \texttt{expl3} programming environment. The \texttt{expl3} modules set up an experimental naming scheme for \LaTeX\ commands, which allow the \LaTeX\ programmer to systematically name functions and variables, and specify the argument types of functions.

The \TeX\ and \epsilon-TeX\ primitives are all given a new name according to these conventions. However, in the main direct use of the primitives is not required or encouraged: the \texttt{expl3} modules define an independent low-level \LaTeX3\ programming language.

At present, the \texttt{expl3} modules are designed to be loaded on top of \LaTeX2\$\epsilon$. In time, a \LaTeX3\ format will be produced based on this code. This allows the code to be used in \LaTeX2\$\epsilon$ packages now while a stand-alone \LaTeX3\ is developed.

While \texttt{expl3} is still experimental, the bundle is now regarded as broadly stable. The syntax conventions and functions provided are now ready for wider use. There may still be changes to some functions, but these will be minor when compared to the scope of \texttt{expl3}.

New modules will be added to the distributed version of \texttt{expl3} as they reach maturity.

\*E-mail: latex-team@latex-project.org
**Contents**

I  Introduction to expl3 and this document  
   1  Naming functions and variables  
      1.1 Terminological inexactitude  
   2  Documentation conventions  
   3  Formal language conventions which apply generally  
   4  \TeX\ concepts not supported by \LaTeX3  

II  The \l3bootstrap\ package: Bootstrap code  
    1  Using the \LaTeX3\ modules  

III  The \l3names\ package: Namespace for primitives  
    1  Setting up the \LaTeX3\ programming language  

IV  The \l3basics\ package: Basic definitions  
    1  No operation functions  
    2  Grouping material  
    3  Control sequences and functions  
       3.1 Defining functions  
       3.2 Defining new functions using parameter text  
       3.3 Defining new functions using the signature  
       3.4 Copying control sequences  
       3.5 Deleting control sequences  
       3.6 Showing control sequences  
       3.7 Converting to and from control sequences  
    4  Analysing control sequences  
    5  Using or removing tokens and arguments  
       5.1 Selecting tokens from delimited arguments  
    6  Predicates and conditionals  
       6.1 Tests on control sequences  
       6.2 Primitive conditionals  
    7  Starting a paragraph  
       7.1 Debugging support  


V  The l3expan package: Argument expansion 25
1  Defining new variants 25
2  Methods for defining variants 26
3  Introducing the variants 27
4  Manipulating the first argument 29
5  Manipulating two arguments 31
6  Manipulating three arguments 32
7  Unbraced expansion 33
8  Preventing expansion 33
9  Controlled expansion 35
10 Internal functions 37

VI The l3tl package: Token lists 38
1  Creating and initialising token list variables 38
2  Adding data to token list variables 39
3  Modifying token list variables 40
4  Reassigning token list category codes 40
5  Token list conditionals 41
6  Mapping to token lists 43
7  Using token lists 46
8  Working with the content of token lists 46
9  The first token from a token list 48
10 Using a single item 51
11 Viewing token lists 53
12 Constant token lists 53
13 Scratch token lists 54

VII The l3str package: Strings 55
1 Building strings 55
2 Adding data to string variables 56
3 Modifying string variables 57
4 String conditionals 58
5 Mapping to strings 59
6 Working with the content of strings 61
7 String manipulation 64
8 Viewing strings 65
9 Constant token lists 66
10 Scratch strings 66

VIII The l3str-convert package: string encoding conversions 67
1 Encoding and escaping schemes 67
2 Conversion functions 67
3 Creating 8-bit mappings 69
4 Possibilities, and things to do 69

IX The l3quark package: Quarks 70
1 Quarks 70
2 Defining quarks 70
3 Quark tests 71
4 Recursion 71
5 An example of recursion with quarks 72
6 Scan marks 73

X The l3seq package: Sequences and stacks 75
1 Creating and initialising sequences 75
2 Appending data to sequences 76
3 Recovering items from sequences 76
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Recovering values from sequences with branching</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>Modifying sequences</td>
<td>79</td>
</tr>
<tr>
<td>6</td>
<td>Sequence conditionals</td>
<td>80</td>
</tr>
<tr>
<td>7</td>
<td>Mapping to sequences</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>Using the content of sequences directly</td>
<td>82</td>
</tr>
<tr>
<td>9</td>
<td>Sequences as stacks</td>
<td>83</td>
</tr>
<tr>
<td>10</td>
<td>Sequences as sets</td>
<td>84</td>
</tr>
<tr>
<td>11</td>
<td>Constant and scratch sequences</td>
<td>85</td>
</tr>
<tr>
<td>12</td>
<td>Viewing sequences</td>
<td>86</td>
</tr>
<tr>
<td>XI</td>
<td>The \texttt{l3int} package: Integers</td>
<td>87</td>
</tr>
<tr>
<td>1</td>
<td>Integer expressions</td>
<td>88</td>
</tr>
<tr>
<td>2</td>
<td>Creating and initialising integers</td>
<td>89</td>
</tr>
<tr>
<td>3</td>
<td>Setting and incrementing integers</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>Using integers</td>
<td>91</td>
</tr>
<tr>
<td>5</td>
<td>Integer expression conditionals</td>
<td>91</td>
</tr>
<tr>
<td>6</td>
<td>Integer expression loops</td>
<td>93</td>
</tr>
<tr>
<td>7</td>
<td>Integer step functions</td>
<td>95</td>
</tr>
<tr>
<td>8</td>
<td>Formatting integers</td>
<td>96</td>
</tr>
<tr>
<td>9</td>
<td>Converting from other formats to integers</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>Random integers</td>
<td>98</td>
</tr>
<tr>
<td>11</td>
<td>Viewing integers</td>
<td>99</td>
</tr>
<tr>
<td>12</td>
<td>Constant integers</td>
<td>99</td>
</tr>
<tr>
<td>13</td>
<td>Scratch integers</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>13.1 Direct number expansion</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>Primitive conditionals</td>
<td>100</td>
</tr>
<tr>
<td>XII</td>
<td>The \texttt{l3flag} package: Expandable flags</td>
<td>102</td>
</tr>
<tr>
<td>1</td>
<td>Setting up flags</td>
<td>102</td>
</tr>
</tbody>
</table>
2 Expandable flag commands 103

XIII The \texttt{l3prg} package: Control structures 104
1 Defining a set of conditional functions 104
2 The boolean data type 106
3 Boolean expressions 108
4 Logical loops 110
5 Producing multiple copies 111
6 Detecting \TeX’s mode 111
7 Primitive conditionals 112
8 Nestable recursions and mappings 112
8.1 Simple mappings ................................................................ 112
9 Internal programming functions 113

XIV The \texttt{l3sys} package: System/runtime functions 114
1 The name of the job 114
2 Date and time 114
3 Engine 114
4 Output format 115
5 Platform 115
6 Random numbers 115
7 Access to the shell 116
7.1 Loading configuration data ..................................................... 117
7.2 Final settings ................................................................. 117

XV The \texttt{l3clist} package: Comma separated lists 118
1 Creating and initialising comma lists 118
2 Adding data to comma lists 120
3 Modifying comma lists 120
4 Comma list conditionals 122
1 Creating new messages 150
2 Contextual information for messages 151
3 Issuing messages 152
4 Redirecting messages 154

XIX The l3file package: File and I/O operations 156
1 Input–output stream management 156
  1.1 Reading from files 157
  1.2 Writing to files 160
  1.3 Wrapping lines in output 162
  1.4 Constant input–output streams, and variables 163
  1.5 Primitive conditionals 163
2 File operation functions 163

XX The l3skip package: Dimensions and skips 168
1 Creating and initialising dim variables 168
2 Setting dim variables 169
3 Utilities for dimension calculations 169
4 Dimension expression conditionals 170
5 Dimension expression loops 172
6 Dimension step functions 173
7 Using dim expressions and variables 174
8 Viewing dim variables 175
9 Constant dimensions 176
10 Scratch dimensions 176
11 Creating and initialising skip variables 176
12 Setting skip variables 177
13 Skip expression conditionals 178
14 Using skip expressions and variables 178
15 Viewing skip variables 178
16  Constant skips 179
17  Scratch skips 179
18  Inserting skips into the output 179
19  Creating and initialising \texttt{muskip} variables 180
20  Setting \texttt{muskip} variables 180
21  Using \texttt{muskip} expressions and variables 181
22  Viewing \texttt{muskip} variables 181
23  Constant muskips 182
24  Scratch muskips 182
25  Primitive conditional 182

XXI  The \texttt{l3keys} package: Key–value interfaces 183
1  Creating keys 184
2  Sub-dividing keys 188
3  Choice and multiple choice keys 189
4  Setting keys 191
5  Handling of unknown keys 191
6  Selective key setting 192
7  Utility functions for keys 193
8  Low-level interface for parsing key–val lists 194

XXII  The \texttt{l3intarray} package: fast global integer arrays 196
1  \texttt{l3intarray} documentation 196
   1.1  Implementation notes ......................................... 197

XXIII  The \texttt{l3fp} package: Floating points 198
1  Creating and initialising floating point variables 199
2  Setting floating point variables 200
3  Using floating points 200

ix
<table>
<thead>
<tr>
<th></th>
<th>Creating and initialising boxes</th>
<th>235</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Using boxes</td>
<td>235</td>
</tr>
<tr>
<td>3</td>
<td>Measuring and setting box dimensions</td>
<td>236</td>
</tr>
<tr>
<td>4</td>
<td>Box conditionals</td>
<td>237</td>
</tr>
<tr>
<td>5</td>
<td>The last box inserted</td>
<td>237</td>
</tr>
<tr>
<td>6</td>
<td>Constant boxes</td>
<td>237</td>
</tr>
<tr>
<td>7</td>
<td>Scratch boxes</td>
<td>238</td>
</tr>
<tr>
<td>8</td>
<td>Viewing box contents</td>
<td>238</td>
</tr>
<tr>
<td>9</td>
<td>Boxes and color</td>
<td>238</td>
</tr>
<tr>
<td>10</td>
<td>Horizontal mode boxes</td>
<td>238</td>
</tr>
<tr>
<td>11</td>
<td>Vertical mode boxes</td>
<td>240</td>
</tr>
<tr>
<td>12</td>
<td>Using boxes efficiently</td>
<td>241</td>
</tr>
<tr>
<td>13</td>
<td>Affine transformations</td>
<td>242</td>
</tr>
<tr>
<td>14</td>
<td>Primitive box conditionals</td>
<td>245</td>
</tr>
</tbody>
</table>

XXIX  **The l3coffins package: Coffin code layer** 246

<table>
<thead>
<tr>
<th></th>
<th>Creating and initialising coffins</th>
<th>246</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Setting coffin content and poles</td>
<td>246</td>
</tr>
<tr>
<td>3</td>
<td>Coffin affine transformations</td>
<td>248</td>
</tr>
<tr>
<td>4</td>
<td>Joining and using coffins</td>
<td>248</td>
</tr>
<tr>
<td>5</td>
<td>Measuring coffins</td>
<td>249</td>
</tr>
<tr>
<td>6</td>
<td>Coffin diagnostics</td>
<td>249</td>
</tr>
<tr>
<td>7</td>
<td>Constants and variables</td>
<td>250</td>
</tr>
</tbody>
</table>

XXX  **The l3color-base package: Color support** 251

|   | Color in boxes                   | 251 |

XXXI  **The l3luatex package: LuaTeX-specific functions** 252

|   | Breaking out to Lua              | 252 |
Part I
Introduction to expl3 and this document

This document is intended to act as a comprehensive reference manual for the expl3 language. A general guide to the \LaTeX3 programming language is found in expl3.pdf.

1 Naming functions and variables

\LaTeX3 does not use @ as a “letter” for defining internal macros. Instead, the symbols _ and : are used in internal macro names to provide structure. The name of each function is divided into logical units using _, while : separates the name of the function from the argument specifier (“arg-spec”). This describes the arguments expected by the function. In most cases, each argument is represented by a single letter. The complete list of arg-spec letters for a function is referred to as the signature of the function.

Each function name starts with the module to which it belongs. Thus apart from a small number of very basic functions, all expl3 function names contain at least one underscore to divide the module name from the descriptive name of the function. For example, all functions concerned with comma lists are in module cist and begin \clist_.

Every function must include an argument specifier. For functions which take no arguments, this will be blank and the function name will end :. Most functions take one or more arguments, and use the following argument specifiers:

N and n These mean no manipulation, of a single token for N and of a set of tokens given in braces for n. Both pass the argument through exactly as given. Usually, if you use a single token for an n argument, all will be well.

c This means csname, and indicates that the argument will be turned into a csname before being used. So \foo:c {ArgumentOne} will act in the same way as \foo:N ArgumentOne.

V and v These mean value of variable. The V and v specifiers are used to get the content of a variable without needing to worry about the underlying \TeX structure containing the data. A V argument will be a single token (similar to N), for example \foo:V MyVariable; on the other hand, using v a csname is constructed first, and then the value is recovered, for example \foo:v {MyVariable}.

o This means expansion once. In general, the V and v specifiers are favoured over o for recovering stored information. However, o is useful for correctly processing information with delimited arguments.

x The x specifier stands for exhaustive expansion: every token in the argument is fully expanded until only unexpandable ones remain. The \TeX \edef primitive carries out this type of expansion. Functions which feature an x-type argument are not expandable.

e The e specifier is in many respects identical to x, but with a very different implementation. Functions which feature an e-type argument may be expandable. The drawback is that e is extremely slow (often more than 200 times slower) in older engines, more precisely in non-Lua\TeX engines older than 2019.
The \texttt{f} specifier stands for \textit{full expansion}, and in contrast to \texttt{x} stops at the first non-expandable token (reading the argument from left to right) without trying to expand it. If this token is a \textit{space token}, it is gobbled, and thus won’t be part of the resulting argument. For example, when setting a token list variable (a macro used for storage), the sequence

\begin{verbatim}
\tl_set:Nn \l_mya_tl { A }
\tl_set:Nn \l_myb_tl { B }
\tl_set:Nf \l_mya_tl { \l_mya_tl \l_myb_tl }
\end{verbatim}

will leave \texttt{\l_mya_tl} with the content \texttt{A\l_myb_tl}, as \texttt{A} cannot be expanded and so terminates expansion before \texttt{\l_myb_tl} is considered.

\textbf{T and F} For logic tests, there are the branch specifiers \texttt{T} (\textit{true}) and \texttt{F} (\textit{false}). Both specifiers treat the input in the same way as \texttt{n} (no change), but make the logic much easier to see.

\textbf{p} The letter \texttt{p} indicates \LaTeX\ \textit{parameters}. Normally this will be used for delimited functions as \texttt{expl3} provides better methods for creating simple sequential arguments.

\textbf{w} Finally, there is the \texttt{w} specifier for \textit{weird} arguments. This covers everything else, but mainly applies to delimited values (where the argument must be terminated by some specified string).

\textbf{D} The \texttt{D} specifier means \textit{do not use}. All of the \LaTeX\ primitives are initially \texttt{\let} to a \texttt{D} name, and some are then given a second name. Only the kernel team should use anything with a \texttt{D} specifier!

Notice that the argument specifier describes how the argument is processed prior to being passed to the underlying function. For example, \texttt{\textbackslash foo:c} will take its argument, convert it to a control sequence and pass it to \texttt{\textbackslash foo:N}.

Variables are named in a similar manner to functions, but begin with a single letter to define the type of variable:

- \texttt{c} Constant: global parameters whose value should not be changed.
- \texttt{g} Parameters whose value should only be set globally.
- \texttt{l} Parameters whose value should only be set locally.

Each variable name is then build up in a similar way to that of a function, typically starting with the module\footnote{\textit{The module names are not used in case of generic scratch registers defined in the data type modules, e.g., the \texttt{int} module contains some scratch variables called \texttt{\l_tmpa_int}, \texttt{\l_tmpb_int}, and so on. In such a case adding the module name up front to denote the module and in the back to indicate the type, as in \texttt{\l_int\_tmpa\_int} would be very unreadable.}} name and then a descriptive part. Variables end with a short identifier to show the variable type:

- \texttt{clist} Comma separated list.
- \texttt{dim} “Rigid” lengths.
- \texttt{fp} Floating-point values;
- \texttt{int} Integer-valued count register.
“Rubber” lengths for use in mathematics.

“Sequence”: a data-type used to implement lists (with access at both ends) and stacks.

“Rubber” lengths.

String variables: contain character data.

Token list variables: placeholder for a token list.

Applying \texttt{V}-type or \texttt{v}-type expansion to variables of one of the above types is supported, while it is not supported for the following variable types:

Either true or false.

Box register.

A “box with handles” — a higher-level data type for carrying out box alignment operations.

Integer that can be incremented expandably.

Fixed-size array of floating point values.

Fixed-size array of integers.

An input or output stream, for reading from or writing to, respectively.

Property list: analogue of dictionary or associative arrays in other languages.

Regular expression.

A word of warning. In this document, and others referring to the \texttt{expl3} programming modules, we often refer to “variables” and “functions” as if they were actual constructs from a real programming language. In truth, \TeX{} is a macro processor, and functions are simply macros that may or may not take arguments and expand to their replacement text. Many of the common variables are also macros, and if placed into the input stream will simply expand to their definition as well — a “function” with no arguments and a “token list variable” are almost the same.\footnote{\TeX{}nically, functions with no arguments are \texttt{\long} while token list variables are not.} On the other hand, some “variables” are actually registers that must be initialised and their values set and retrieved with specific functions.

The conventions of the \texttt{expl3} code are designed to clearly separate the ideas of “macros that contain data” and “macros that contain code”, and a consistent wrapper is applied to all forms of “data” whether they be macros or actually registers. This means that sometimes we will use phrases like “the function returns a value”, when actually we just mean “the macro expands to something”. Similarly, the term “execute” might be used in place of “expand” or it might refer to the more specific case of “processing in \TeX{}’s stomach” (if you are familiar with the \TeX{}book parlance).

If in doubt, please ask; chances are we’ve been hasty in writing certain definitions and need to be told to tighten up our terminology.

1.1 Terminological inexactitude
2 Documentation conventions

This document is typeset with the experimental l3doc class; several conventions are used to help describe the features of the code. A number of conventions are used here to make the documentation clearer.

Each group of related functions is given in a box. For a function with a “user” name, this might read:

\ExplSyntaxOn ...
\ExplSyntaxOff

The textual description of how the function works would appear here. The syntax of the function is shown in mono-spaced text to the right of the box. In this example, the function takes no arguments and so the name of the function is simply reprinted.

For programming functions, which use _ and : in their name there are a few additional conventions: If two related functions are given with identical names but different argument specifiers, these are termed variants of each other, and the latter functions are printed in grey to show this more clearly. They will carry out the same function but will take different types of argument:

\seq_new:N \seq_new:c

When a number of variants are described, the arguments are usually illustrated only for the base function. Here, (sequence) indicates that \seq_new:N expects the name of a sequence. From the argument specifier, \seq_new:c also expects a sequence name, but as a name rather than as a control sequence. Each argument given in the illustration should be described in the following text.

Fully expandable functions Some functions are fully expandable, which allows them to be used within an x-type or e-type argument (in plain \TeX terms, inside an \edef or \expanded), as well as within an f-type argument. These fully expandable functions are indicated in the documentation by a star:

\cs_to_str:N ★ \cs_to_str:N (cs)

As with other functions, some text should follow which explains how the function works. Usually, only the star will indicate that the function is expandable. In this case, the function expects a (cs), shorthand for a (control sequence).

Restricted expandable functions A few functions are fully expandable but cannot be fully expanded within an f-type argument. In this case a hollow star is used to indicate this:

\seq_map_function:NN ★ \seq_map_function:NN (seq) (function)

Conditional functions Conditional (if) functions are normally defined in three variants, with T, F and TF argument specifiers. This allows them to be used for different “true”/“false” branches, depending on which outcome the conditional is being used to test. To indicate this without repetition, this information is given in a shortened form:
The underlining and italic of \textit{TF} indicates that three functions are available:

- \texttt{\textbackslash sys\_if\_engine\_xetex:T}
- \texttt{\textbackslash sys\_if\_engine\_xetex:F}
- \texttt{\textbackslash sys\_if\_engine\_xetex:TF}

Usually, the illustration will use the \texttt{TF} variant, and so both \texttt{(true code)} and \texttt{(false code)} will be shown. The two variant forms \texttt{T} and \texttt{F} take only \texttt{(true code)} and \texttt{(false code)}, respectively. Here, the star also shows that this function is expandable. With some minor exceptions, \textit{all} conditional functions in the \texttt{expl3} modules should be defined in this way.

Variables, constants and so on are described in a similar manner:

\begin{verbatim}
\l_tmpa_tl
\end{verbatim}

A short piece of text will describe the variable: there is no syntax illustration in this case.

In some cases, the function is similar to one in \LaTeX{}2ε or plain \TeX{}. In these cases, the text will include an extra \textit{“\TeX{}hackers note”} section:

\begin{verbatim}
\token_to_str:N
\end{verbatim}

The normal description text.

\textbf{\TeX{}hackers note}: Detail for the experienced \TeX{} or \LaTeX{}2ε programmer. In this case, it would point out that this function is the \TeX{} primitive \texttt{\string}.

Changes to behaviour: When new functions are added to \texttt{expl3}, the date of first inclusion is given in the documentation. Where the documented behaviour of a function changes after it is first introduced, the date of the update will also be given. This means that the programmer can be sure that any release of \texttt{expl3} after the date given will contain the function of interest with expected behaviour as described. Note that changes to code internals, including bug fixes, are not recorded in this way unless they impact on the expected behaviour.

3 Formal language conventions which apply generally

As this is a formal reference guide for \LaTeX{}3 programming, the descriptions of functions are intended to be reasonably “complete”. However, there is also a need to avoid repetition. Formal ideas which apply to general classes of function are therefore summarised here.

For tests which have a \texttt{TF} argument specification, the test if evaluated to give a logically \texttt{TRUE} or \texttt{FALSE} result. Depending on this result, either the \texttt{(true code)} or the \texttt{(false code)} will be left in the input stream. In the case where the test is expandable, and a predicate (_p) variant is available, the logical value determined by the test is left in the input stream: this will typically be part of a larger logical construct.
4  \TeX \text{ concepts not supported by \LaTeX{}3}

The \TeX{} concept of an “\texttt{\textbackslash outer}” macro is \textit{not supported} at all by \LaTeX{}3. As such, the functions provided here may break when used on top of \LaTeX{}2e if \texttt{\textbackslash outer} tokens are used in the arguments.
Part II

The l3bootstrap package

Bootstrap code

1 Using the \TeX3 modules

The modules documented in source3 are designed to be used on top of \TeX2ε and are loaded all as one with the usual \usepackage{expl3} or \RequirePackage{expl3} instructions. These modules will also form the basis of the \TeXX3 format, but work in this area is incomplete and not included in this documentation at present.

As the modules use a coding syntax different from standard \TeX2ε it provides a few functions for setting it up.

\ExplSyntaxOn
\ExplSyntaxOff

The \ExplSyntaxOn function switches to a category code régime in which spaces are ignored and in which the colon (:) and underscore (_) are treated as “letters”, thus allowing access to the names of code functions and variables. Within this environment, ~ is used to input a space. The \ExplSyntaxOff reverts to the document category code régime.

\ExplSyntaxOn
\ExplSyntaxOff

Updated: 2011-08-13

\RequirePackage{expl3}
\ProvidesExplPackage {\package} {\date} {\version} {\description}

These functions act broadly in the same way as the corresponding \TeX2ε kernel functions \ProvidesPackage, \ProvidesClass and \ProvidesFile. However, they also implicitly switch \ExplSyntaxOn for the remainder of the code with the file. At the end of the file, \ExplSyntaxOff will be called to reverse this. (This is the same concept as \TeX2ε provides in turning on \makeatletter within package and class code.) The \date should be given in the format \year/\month/\day. If the \version is given then it will be prefixed with v in the package identifier line.

\RequirePackage{expl3}
\ProvidesExplPackage {\ExplFileName} {\ExplFileDate} {\ExplFileVersion} {\ExplFileDescription}

Updated: 2017-03-19

\GetIdInfo $Id: \svn info field $ {\description}

Extracts all information from a SVN field. Spaces are not ignored in these fields. The information pieces are stored in separate control sequences with \ExplFileName for the part of the file name leading up to the period, \ExplFileDate for date, \ExplFileVersion for version and \ExplFileDescription for the description.

To summarize: Every single package using this syntax should identify itself using one of the above methods. Special care is taken so that every package or class file loaded with \RequirePackage or similar are loaded with usual \TeX2ε category codes and the \TeXX3 category code scheme is reloaded when needed afterwards. See implementation for details. If you use the \GetIdInfo command you can use the information when loading a package with

\ProvidesExplPackage{\ExplFileName}
{\ExplFileDate}{\ExplFileVersion}{\ExplFileDescription}
Part III

The l3names package
Namespace for primitives

1 Setting up the \LaTeX3 programming language

This module is at the core of the \LaTeX3 programming language. It performs the following tasks:

- defines new names for all \TeX primitives;
- switches to the category code régime for programming;
- provides support settings for building the code as a \TeX format.

This module is entirely dedicated to primitives, which should not be used directly within \LaTeX3 code (outside of “kernel-level” code). As such, the primitives are not documented here: The \TeXbook, \TeX by Topic and the manuals for pdf\TeX, Xe\TeX, Lua\TeX, \mu\TeX and up\TeX should be consulted for details of the primitives. These are named \texttt{\tex\langle name\rangle}; typically based on the primitive’s \langle name\rangle in pdf\TeX and omitting a leading pdf when the primitive is not related to pdf output.
Part IV
The \texttt{l3basics} package

Basic definitions

As the name suggest this package holds some basic definitions which are needed by most or all other packages in this set.

Here we describe those functions that are used all over the place. With that we mean functions dealing with the construction and testing of control sequences. Furthermore the basic parts of conditional processing are covered; conditional processing dealing with specific data types is described in the modules specific for the respective data types.

1 No operation functions

\begin{verbatim}
\prg_do_nothing: * \prg_do_nothing:
\end{verbatim}
An expandable function which does nothing at all: leaves nothing in the input stream after a single expansion.

\begin{verbatim}
\scan_stop: \scan_stop:
\end{verbatim}
A non-expandable function which does nothing. Does not vanish on expansion but produces no typeset output.

2 Grouping material

\begin{verbatim}
\group_begin: \group_begin:
\group_end: \group_end:
\end{verbatim}
These functions begin and end a group for definition purposes. Assignments are local to groups unless carried out in a global manner. (A small number of exceptions to this rule will be noted as necessary elsewhere in this document.) Each \texttt{\group_begin}: must be matched by a \texttt{\group_end}:; although this does not have to occur within the same function. Indeed, it is often necessary to start a group within one function and finish it within another, for example when seeking to use non-standard category codes.

\begin{verbatim}
\group_insert_after:N \group_insert_after:N \token\)
\end{verbatim}
Adds \texttt{\token} to the list of \texttt{\tokens} to be inserted when the current group level ends. The list of \texttt{\tokens} to be inserted is empty at the beginning of a group: multiple applications of \texttt{\group_insert_after:N} may be used to build the inserted list one \texttt{\token} at a time. The current group level may be closed by a \texttt{\group_end}: function or by a token with category code 2 (close-group), namely a } if standard category codes apply.
3 Control sequences and functions

As \TeX is a macro language, creating new functions means creating macros. At point of use, a function is replaced by the replacement text (“code”) in which each parameter in the code (#1, #2, etc.) is replaced the appropriate arguments absorbed by the function. In the following, (code) is therefore used as a shorthand for “replacement text”.

Functions which are not “protected” are fully expanded inside an x expansion. In contrast, “protected” functions are not expanded within x expansions.

3.1 Defining functions

Functions can be created with no requirement that they are declared first (in contrast to variables, which must always be declared). Declaring a function before setting up the code means that the name chosen is checked and an error raised if it is already in use. The name of a function can be checked at the point of definition using the \cs_new... functions: this is recommended for all functions which are defined for the first time.

There are three ways to define new functions. All classes define a function to expand to the substitution text. Within the substitution text the actual parameters are substituted for the formal parameters (#1, #2, ...).

new Create a new function with the new scope, such as \cs_new:Npn. The definition is global and results in an error if it is already defined.

set Create a new function with the set scope, such as \cs_set:Npn. The definition is restricted to the current \TeX group and does not result in an error if the function is already defined.

gset Create a new function with the gset scope, such as \cs_gset:Npn. The definition is global and does not result in an error if the function is already defined.

Within each set of scope there are different ways to define a function. The differences depend on restrictions on the actual parameters and the expandability of the resulting function.

nopar Create a new function with the nopar restriction, such as \cs_set_nopar:Npn. The parameter may not contain \par tokens.

protected Create a new function with the protected restriction, such as \cs_set_-_protected:Npn. The parameter may contain \par tokens but the function will not expand within an x-type or e-type expansion.

Finally, the functions in Subsections 3.2 and 3.3 are primarily meant to define base functions only. Base functions can only have the following argument specifiers:

\textbf{N} and \textbf{n} No manipulation.

\textbf{T} and \textbf{F} Functionally equivalent to \textbf{n} (you are actually encouraged to use the family of \prg_new_conditional: functions described in Section 1).

\textbf{p} and \textbf{w} These are special cases.

The \cs_new: functions below (and friends) do not stop you from using other argument specifiers in your function names, but they do not handle expansion for you. You should define the base function and then use \cs_generate_variant:Nn to generate custom variants as described in Section 2.
3.2 Defining new functions using parameter text

\cs_new: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Creates \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. The definition is global and an error results if the \langle function \rangle is already defined.

\cs_new nopar: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Creates \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. When the \langle function \rangle is used the \langle parameters \rangle absorbed cannot contain \par tokens. The definition is global and an error results if the \langle function \rangle is already defined.

\cs_new protected: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Creates \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. The \langle function \rangle will not expand within an x-type argument. The definition is global and an error results if the \langle function \rangle is already defined.

\cs_new protected nopar: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Creates \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. The definition is global and an error results if the \langle function \rangle is already defined.

\cs_set: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. The assignment of a meaning to the \langle function \rangle is restricted to the current T\LaTeX{} group level.

\cs_set nopar: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. When the \langle function \rangle is used the \langle parameters \rangle absorbed cannot contain \par tokens. The assignment of a meaning to the \langle function \rangle is restricted to the current T\LaTeX{} group level.

\cs_set protected: Npn \langle function \rangle \langle parameters \rangle \{\langle code \rangle\}

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the \langle parameters \rangle (#1, #2, etc.) will be replaced by those absorbed by the function. The assignment of a meaning to the \langle function \rangle is restricted to the current T\LaTeX{} group level. The \langle function \rangle will not expand within an x-type or e-type argument.
Sets \texttt{(function)} to expand to \texttt{(code)} as replacement text. Within the \texttt{(code)}, the \texttt{(parameters)} (#1, #2, etc.) will be replaced by those absorbed by the function. When the \texttt{(function)} is used the \texttt{(parameters)} absorbed cannot contain \texttt{\par} tokens. The assignment of a meaning to the \texttt{(function)} is restricted to the current \TeX{} group level. The \texttt{(function)} will not expand within an x-type or e-type argument.

Globally sets \texttt{(function)} to expand to \texttt{(code)} as replacement text. Within the \texttt{(code)}, the \texttt{(parameters)} (#1, #2, etc.) will be replaced by those absorbed by the function. The assignment of a meaning to the \texttt{(function)} is not restricted to the current \TeX{} group level: the assignment is global.

Globally sets \texttt{(function)} to expand to \texttt{(code)} as replacement text. Within the \texttt{(code)}, the \texttt{(parameters)} (#1, #2, etc.) will be replaced by those absorbed by the function. When the \texttt{(function)} is used the \texttt{(parameters)} absorbed cannot contain \texttt{\par} tokens. The assignment of a meaning to the \texttt{(function)} is not restricted to the current \TeX{} group level: the assignment is global. The \texttt{(function)} will not expand within an x-type or e-type argument.

Globally sets \texttt{(function)} to expand to \texttt{(code)} as replacement text. Within the \texttt{(code)}, the \texttt{(parameters)} (#1, #2, etc.) will be replaced by those absorbed by the function. When the \texttt{(function)} is used the \texttt{(parameters)} absorbed cannot contain \texttt{\par} tokens. The assignment of a meaning to the \texttt{(function)} is not restricted to the current \TeX{} group level: the assignment is global. The \texttt{(function)} will not expand within an x-type argument.

\subsection{3.3 Defining new functions using the signature}

Creates \texttt{(function)} to expand to \texttt{(code)} as replacement text. Within the \texttt{(code)}, the number of \texttt{(parameters)} is detected automatically from the function signature. These \texttt{(parameters)} (#1, #2, etc.) will be replaced by those absorbed by the function. The definition is global and an error results if the \texttt{(function)} is already defined.
\cs_set_protected:Nn
\cs_set_protected:(cn|Nx|cx)
\cs_new_protected:Nn (function) \{\code\}

Creates \textit{(function)} to expand to \textit{\{code\}} as replacement text. Within the \textit{\{code\}}, the number of \textit{\{parameters\}} is detected automatically from the function signature. These \textit{\{parameters\}} \((#1, #2, \text{etc.})\) will be replaced by those absorbed by the function. When the \textit{\{function\}} is used the \textit{\{parameters\}} absorbed cannot contain \texttt{\par} tokens. The definition is global and an error results if the \textit{\{function\}} is already defined.

\cs_set_protected:Nn
\cs_set_protected:(cn|Nx|cx)
\cs_new_protected:nopar:Nn (function) \{\code\}

Creates \textit{(function)} to expand to \textit{\{code\}} as replacement text. Within the \textit{\{code\}}, the number of \textit{\{parameters\}} is detected automatically from the function signature. These \textit{\{parameters\}} \((#1, #2, \text{etc.})\) will be replaced by those absorbed by the function. The \textit{\{function\}} will not expand within an x-type argument. The definition is global and an error results if the \textit{\{function\}} is already defined.

\cs_set_protected:nopar:Nn
\cs_set_protected:nopar:(cn|Nx|cx)
\cs_new_protected:nopar:Nn (function) \{\code\}

Creates \textit{(function)} to expand to \textit{\{code\}} as replacement text. Within the \textit{\{code\}}, the number of \textit{\{parameters\}} is detected automatically from the function signature. These \textit{\{parameters\}} \((#1, #2, \text{etc.})\) will be replaced by those absorbed by the function. The \textit{\{function\}} will not expand within an x-type or e-type argument. The definition is global and an error results if the \textit{\{function\}} is already defined.

\cs_set:Nn
\cs_set:(cn|Nx|cx)
\cs_set:Nn (function) \{\code\}

Sets \textit{(function)} to expand to \textit{\{code\}} as replacement text. Within the \textit{\{code\}}, the number of \textit{\{parameters\}} is detected automatically from the function signature. These \textit{\{parameters\}} \((#1, #2, \text{etc.})\) will be replaced by those absorbed by the function. The assignment of a meaning to the \textit{\{function\}} is restricted to the current \TeX{} group level.

\cs_set:nopar:Nn
\cs_set:nopar:(cn|Nx|cx)
\cs_set:nopar:Nn (function) \{\code\}

Sets \textit{(function)} to expand to \textit{\{code\}} as replacement text. Within the \textit{\{code\}}, the number of \textit{\{parameters\}} is detected automatically from the function signature. These \textit{\{parameters\}} \((#1, #2, \text{etc.})\) will be replaced by those absorbed by the function. When the \textit{\{function\}} is used the \textit{\{parameters\}} absorbed cannot contain \texttt{\par} tokens. The assignment of a meaning to the \textit{\{function\}} is restricted to the current \TeX{} group level.

\cs_set:nopar:prot:Nn
\cs_set:nopar:prot:(cn|Nx|cx)
\cs_set:prot:Nn (function) \{\code\}

Sets \textit{(function)} to expand to \textit{\{code\}} as replacement text. Within the \textit{\{code\}}, the number of \textit{\{parameters\}} is detected automatically from the function signature. These \textit{\{parameters\}} \((#1, #2, \text{etc.})\) will be replaced by those absorbed by the function. The \textit{\{function\}} will not expand within an x-type argument. The assignment of a meaning to the \textit{\{function\}} is restricted to the current \TeX{} group level.
\cs_set_protected_nopar:Nn \cs_set_protected_nopar:(cn|Nx|cx) \cs_set_protected_nopar:Nn (function) \{code\}

Sets (function) to expand to (code) as replacement text. Within the (code), the number of (parameters) is detected automatically from the function signature. These (parameters) (#1, #2, etc.) will be replaced by those absorbed by the function. When the (function) is used the (parameters) absorbed cannot contain \par tokens. The (function) will not expand within an x-type or e-type argument. The assignment of a meaning to the (function) is restricted to the current \TeX{} group level.

\cs_gset:Nn \cs_gset:(cn|Nx|cx) \cs_gset:Nn (function) \{code\}

Sets (function) to expand to (code) as replacement text. Within the (code), the number of (parameters) is detected automatically from the function signature. These (parameters) (#1, #2, etc.) will be replaced by those absorbed by the function. The assignment of a meaning to the (function) is global.

\cs_gset_nopar:Nn \cs_gset_nopar:(cn|Nx|cx) \cs_gset_nopar:Nn (function) \{code\}

Sets (function) to expand to (code) as replacement text. Within the (code), the number of (parameters) is detected automatically from the function signature. These (parameters) (#1, #2, etc.) will be replaced by those absorbed by the function. When the (function) is used the (parameters) absorbed cannot contain \par tokens. The assignment of a meaning to the (function) is global.

\cs_gset_protected:Nn \cs_gset_protected:(cn|Nx|cx) \cs_gset_protected:Nn (function) \{code\}

Sets (function) to expand to (code) as replacement text. Within the (code), the number of (parameters) is detected automatically from the function signature. These (parameters) (#1, #2, etc.) will be replaced by those absorbed by the function. The (function) will not expand within an x-type argument. The assignment of a meaning to the (function) is global.

\cs_gset_protected_nopar:Nn \cs_gset_protected_nopar:(cn|Nx|cx) \cs_gset_protected_nopar:Nn (function) \{code\}

Sets (function) to expand to (code) as replacement text. Within the (code), the number of (parameters) is detected automatically from the function signature. These (parameters) (#1, #2, etc.) will be replaced by those absorbed by the function. When the (function) is used the (parameters) absorbed cannot contain \par tokens. The (function) will not expand within an x-type or e-type argument. The assignment of a meaning to the (function) is global.

\cs_generate_from_arg_count:NNN \cs_generate_from_arg_count:NNN (function) (creator) \{number\} \{code\}

Uses the (creator) function (which should have signature Nnp, for example \cs_new:Nnpn) to define a (function) which takes (number) arguments and has (code) as replacement text. The (number) of arguments is an integer expression, evaluated as detailed for \int_eval:n.
3.4 Copying control sequences

Control sequences (not just functions as defined above) can be set to have the same meaning using the functions described here. Making two control sequences equivalent means that the second control sequence is a copy of the first (rather than a pointer to it). Thus the old and new control sequence are not tied together: changes to one are not reflected in the other.

In the following text “cs” is used as an abbreviation for “control sequence”.

\begin{verbatim}
\cs_new_eq:NN \cs_new_eq:NN \cs_new_eq:NN \cs_new_eq:NN
\cs_new_eq:NN \cs_new_eq:NN \cs_new_eq:NN \cs_new_eq:NN
\cs_set_eq:NN \cs_set_eq:NN \cs_set_eq:NN \cs_set_eq:NN
\cs_set_eq:NN \cs_set_eq:NN \cs_set_eq:NN \cs_set_eq:NN
\cs_gset_eq:NN \cs_gset_eq:NN \cs_gset_eq:NN \cs_gset_eq:NN
\cs_gset_eq:NN \cs_gset_eq:NN \cs_gset_eq:NN \cs_gset_eq:NN
\cs_undefine:N \cs_undefine:N \cs_undefine:N \cs_undefine:N
\cs_undefine:N \cs_undefine:N \cs_undefine:N \cs_undefine:N
\cs_meaning:N \cs_meaning:N \cs_meaning:N \cs_meaning:N
\cs_meaning:N \cs_meaning:N \cs_meaning:N \cs_meaning:N
\cs_meaning:N \cs_meaning:N \cs_meaning:N \cs_meaning:N
\cs_meaning:N \cs_meaning:N \cs_meaning:N \cs_meaning:N
\end{verbatim}

3.5 Deleting control sequences

There are occasions where control sequences need to be deleted. This is handled in a very simple manner.

\begin{verbatim}
\cs_undefine:N \cs_undefine:N \cs_undefine:N \cs_undefine:N
\cs_undefine:N \cs_undefine:N \cs_undefine:N \cs_undefine:N
\end{verbatim}

3.6 Showing control sequences

\begin{verbatim}
\cs_meaning:N \cs_meaning:N \cs_meaning:N \cs_meaning:N
\cs_meaning:N \cs_meaning:N \cs_meaning:N \cs_meaning:N
\end{verbatim}

\TeXhackers note: This is \TeX’s \texttt{\meaning} primitive. For tokens that are not control sequences, it is more logical to use \texttt{\token_to_meaning:N}. The \texttt{c} variant correctly reports undefined arguments.

Updated: 2011-09-15

Updated: 2011-12-22
\cs_show:N \cs_show:c
Displays the definition of the \textit{control sequence} on the terminal.

\textbf{\TeX{}hackers note:} This is similar to the \TeX{} primitive \texttt{\show}, wrapped to a fixed number of characters per line.

\cs_log:N \cs_log:c
Write the definition of the \textit{control sequence} in the log file. See also \texttt{\cs_show:N} which displays the result in the terminal.

### 3.7 Converting to and from control sequences

\use:c \{\textit{control sequence name}\}\}
Expands the \textit{control sequence name} until only characters remain, and then converts this into a control sequence. This process requires two expansions. As in other \texttt{c}-type arguments the \textit{control sequence name} must, when fully expanded, consist of character tokens, typically a mixture of category code 10 (space), 11 (letter) and 12 (other).

\textbf{\TeX{}hackers note:} Protected macros that appear in a \texttt{c}-type argument are expanded despite being protected; \texttt{\exp_not:n} also has no effect. An internal error occurs if non-characters or active characters remain after full expansion, as the conversion to a control sequence is not possible.

As an example of the \texttt{\use:c} function, both

\begin{verbatim}
\use:c \{ a b c \}
\end{verbatim}

and

\begin{verbatim}
\tl_new:N \l_my_tl
\tl_set:Nn \l_my_tl \{ a b c \}
\use:c \{ \tl_use:N \l_my_tl \}
\end{verbatim}

would be equivalent to

\begin{verbatim}
\abc
\end{verbatim}

after two expansions of \texttt{\use:c}.

\cs_if_exist_use:N \cs_if_exist_use:c \cs_if_exist_use:NTF \cs_if_exist_use:cTF
Tests whether the \textit{control sequence} is currently defined according to the conditional \texttt{\cs_if_exist:NTF} (whether as a function or another control sequence type), and if it is inserts the \textit{control sequence} into the input stream followed by the \textit{true code}. Otherwise the \textit{false code} is used.

\textbf{\TeX{}hackers note:}
\texttt{\textbackslash cs:w} \texttt{\{control sequence name\} \textbackslash cs_end:}

Converts the given \texttt{\{control sequence name\}} into a single control sequence token. This process requires one expansion. The content for \texttt{\{control sequence name\}} may be literal material or from other expandable functions. The \texttt{\{control sequence name\}} must, when fully expanded, consist of character tokens which are not active: typically of category code 10 (space), 11 (letter) or 12 (other), or a mixture of these.

\textbf{\texttt{\textbackslash cs:end:} note:} These are the \TeX\ primitives \texttt{\textbackslash csname} and \texttt{\textbackslash endcsname}.

As an example of the \texttt{\textbackslash cs:w} and \texttt{\textbackslash cs_end:} functions, both

\begin{verbatim}
\texttt{\textbackslash cs:w a b c \textbackslash cs_end:}
\end{verbatim}

and

\begin{verbatim}
\texttt{\tl_new:N \l_my_tl}
\texttt{\tl_set:Nn \l_my_tl \{ a b c \}}
\texttt{\textbackslash cs:w \tl_use:N \l_my_tl \textbackslash cs_end:}
\end{verbatim}

would be equivalent to

\begin{verbatim}
\texttt{\textbackslash abc}
\end{verbatim}

after one expansion of \texttt{\textbackslash cs:w}.

\texttt{\textbackslash cs_to_str:N} \texttt{\{control sequence\}}

Converts the given \texttt{\{control sequence\}} into a series of characters with category code 12 (other), except spaces, of category code 10. The result does \textit{not} include the current escape token, contrarily to \texttt{\textbackslash token_to_str:N}. Full expansion of this function requires exactly 2 expansion steps, and so an \texttt{x}-type or \texttt{e}-type expansion, or two \texttt{o}-type expansions are required to convert the \texttt{\{control sequence\}} to a sequence of characters in the input stream. In most cases, an \texttt{f}-expansion is correct as well, but this loses a space at the start of the result.

4 Analysing control sequences

\texttt{\textbackslash cs_split_function:N} \texttt{\{function\}}

Splits the \texttt{\{function\}} into the \texttt{\{name\}} (i.e. the part before the colon) and the \texttt{\{signature\}} (i.e. after the colon). This information is then placed in the input stream in three parts: the \texttt{\{name\}}, the \texttt{\{signature\}} and a logic token indicating if a colon was found (to differentiate variables from function names). The \texttt{\{name\}} does not include the escape character, and both the \texttt{\{name\}} and \texttt{\{signature\}} are made up of tokens with category code 12 (other).

The next three functions decompose \TeX\ macros into their constituent parts: if the \texttt{\{token\}} passed is not a macro then no decomposition can occur. In the latter case, all three functions leave \texttt{\textbackslash scan_stop:} in the input stream.
If the \langle \text{token} \rangle is a macro, this function leaves the applicable \TeX\ prefixes in input stream as a string of tokens of category code 12 (with spaces having category code 10). Thus for example

\begin{verbatim}
\cs_set:Npn \next:nn #1 #2 { x #1 \textasciitilde y #2 }
\cs_prefix_spec:N \next:nn
\end{verbatim}

leaves \texttt{\long} in the input stream. If the \langle \text{token} \rangle is not a macro then \texttt{\scan_stop:} is left in the input stream.

\texttt{\TeXhackers note:} The prefix can be empty, \texttt{\long}, \texttt{\protected} or \texttt{\protected\long} with backslash replaced by the current escape character.

If the \langle \text{token} \rangle is a macro, this function leaves the primitive \TeX\ argument specification in input stream as a string of character tokens of category code 12 (with spaces having category code 10). Thus for example

\begin{verbatim}
\cs_set:Npn \next:nn #1 #2 { x #1 y #2 }
\cs_argument_spec:N \next:nn
\end{verbatim}

leaves \#1\#2 in the input stream. If the \langle \text{token} \rangle is not a macro then \texttt{\scan_stop:} is left in the input stream.

\texttt{\TeXhackers note:} If the argument specification contains the string ->', then the function produces incorrect results.

If the \langle \text{token} \rangle is a macro, this function leaves the replacement text in input stream as a string of character tokens of category code 12 (with spaces having category code 10). Thus for example

\begin{verbatim}
\cs_set:Npn \next:nn #1 #2 { x #1 \textasciitilde y #2 }
\cs_replacement_spec:N \next:nn
\end{verbatim}

leaves \texttt{x#1\textasciitilde y#2} in the input stream. If the \langle \text{token} \rangle is not a macro then \texttt{\scan_stop:} is left in the input stream.

\texttt{\TeXhackers note:} If the argument specification contains the string ->', then the function produces incorrect results.

5 Using or removing tokens and arguments

Tokens in the input can be read and used or read and discarded. If one or more tokens are wrapped in braces then when absorbing them the outer set is removed. At the same time, the category code of each token is set when the token is read by a function (if it
is read more than once, the category code is determined by the situation in force when
first function absorbs the token).

\use:n \use:nn \use:nnn \use:nnnn

As illustrated, these functions absorb between one and four arguments, as indicated by
the argument specifier. The braces surrounding each argument are removed and the
remaining tokens are left in the input stream. The category code of these tokens is also
fixed by this process (if it has not already been by some other absorption). All of these
functions require only a single expansion to operate, so that one expansion of

\use:nn \{ abc \} \{ \{ def \} \}

results in the input stream containing

abc \{ def \}

\TeXhackers note: The \use:n function is equivalent to \LaTeX\’s \@firstofone.

\use:i:nn \use:ii:nn \use:i:nnn

These functions absorb two arguments from the input stream. The function \use:i:nn
discards the second argument, and leaves the content of the first argument in the input
stream. \use:ii:nn discards the first argument and leaves the content of the second
argument in the input stream. The category code of these tokens is also fixed (if it
has not already been by some other absorption). A single expansion is needed for the
functions to take effect.

\TeXhackers note: These are equivalent to \LaTeX\’s \@firstoftwo and \@secondoftwo.

\use:i:nnn \use:ii:nnn \use:iii:nnn \use:iv:nnnn

These functions absorb three arguments from the input stream. The function \use:i:nnn
discards the second and third arguments, and leaves the content of the first argument in
the input stream. \use:ii:nnn and \use:iii:nnn work similarly, leaving the content of
second or third arguments in the input stream, respectively. The category code of
these tokens is also fixed (if it has not already been by some other absorption). A single
expansion is needed for the functions to take effect.

\TeXhackers note: These functions absorb four arguments from the input stream. The function \use:i:nnnn
discards the second, third and fourth arguments, and leaves the content of the
first argument in the input stream. \use:ii:nnnn, \use:iii:nnnn and \use:iv:nnnn
work similarly, leaving the content of second, third or fourth arguments in the input
stream, respectively. The category code of these tokens is also fixed (if it has not already
been by some other absorption). A single expansion is needed for the functions to take
effect.
This function absorbs three arguments and leaves the content of the first and second in the input stream. The category code of these tokens is also fixed (if it has not already been by some other absorption). A single expansion is needed for the function to take effect. An example:

\use_i_ii:nnn \{ abc \} \{ \{ def \} \} \{ ghi \}

results in the input stream containing

abc \{ def \}

i.e. the outer braces are removed and the third group is removed.

This function absorbs two arguments and leaves the content of the second and first in the input stream. The category code of these tokens is also fixed (if it has not already been by some other absorption). A single expansion is needed for the function to take effect.

\use_ii_i:nn

These functions absorb between one and nine groups from the input stream, leaving nothing on the resulting input stream. These functions work after a single expansion. One or more of the \(n\) arguments may be an unbraced single token (i.e. an \(N\) argument).

\use:x {\{expandable tokens\}}

Fully expands the \{expandable tokens\} and inserts the result into the input stream at the current location. Any hash characters (#) in the argument must be doubled.

\use:e {\{expandable tokens\}}

Fully expands the \{expandable tokens\} in an \(x\)-type manner, but the function remains fully expandable, and parameter character (usually \#) need not be doubled.

\use:ii:nnn

This function absorbs three arguments and leaves the content of the first and second in the input stream. The category code of these tokens is also fixed (if it has not already been by some other absorption). A single expansion is needed for the function to take effect. An example:

\use_i_ii:nnn \{ arg_1 \} \{ arg_2 \} \{ arg_3 \}

\use:ii_i:nn

\use:none:n

\use:none:nn

\use:none:nnn

\use:none:nnnn

\use:none:nnnnn

\use:none:nnnnnn

\use:none:nnnnnnn

\use:none:nnnnnnnn

\use:none:nnnnnnnnn

\use:ii:nn

\use:ii_i:nn

\use:ii:nn

\use:none:nn

\use:none:nnn

\use:none:nnnn

\use:none:nnnnn

\use:none:nnnnnn

\use:none:nnnnnnn

\use:none:nnnnnnnn

\use:none:nnnnnnnnn

\use:e

\use:x

5.1 Selecting tokens from delimited arguments

A different kind of function for selecting tokens from the token stream are those that use delimited arguments.
Absorb the \langle balanced text \rangle from the input stream delimited by the marker given in the function name, leaving nothing in the input stream.

\begin{verbatim}
\use_none_delimit_by_q_nil:w *
\use_none_delimit_by_q_stop:w *
\use_none_delimit_by_q_recursion_stop:w *
\end{verbatim}

Absorb the \langle balanced text \rangle from the input stream delimited by the marker given in the function name, leaving \langle inserted tokens \rangle in the input stream for further processing.

\section{Predicates and conditionals}

\LaTeX{} has three concepts for conditional flow processing:

**Branching conditionals** Functions that carry out a test and then execute, depending on its result, either the code supplied as the \langle true code \rangle or the \langle false code \rangle. These arguments are denoted with T and F, respectively. An example would be

\begin{verbatim}
\cs_if_free:cTF {abc} {\langle true code \rangle} {\langle false code \rangle}
\end{verbatim}

a function that turns the first argument into a control sequence (since it’s marked as c) then checks whether this control sequence is still free and then depending on the result carries out the code in the second argument (true case) or in the third argument (false case).

These type of functions are known as “conditionals”; whenever a TF function is defined it is usually accompanied by T and F functions as well. These are provided for convenience when the branch only needs to go a single way. Package writers are free to choose which types to define but the kernel definitions always provide all three versions.

Important to note is that these branching conditionals with \langle true code \rangle and/or \langle false code \rangle are always defined in a way that the code of the chosen alternative can operate on following tokens in the input stream.

These conditional functions may or may not be fully expandable, but if they are expandable they are accompanied by a “predicate” for the same test as described below.

**Predicates** “Predicates” are functions that return a special type of boolean value which can be tested by the boolean expression parser. All functions of this type are expandable and have names that end with _p in the description part. For example,

\begin{verbatim}
\cs_if_free_p:N
\end{verbatim}

would be a predicate function for the same type of test as the conditional described above. It would return “true” if its argument (a single token denoted by N) is still free for definition. It would be used in constructions like
For each predicate defined, a “branching conditional” also exists that behaves like a conditional described above.

**Primitive conditionals** There is a third variety of conditional, which is the original concept used in plain TeX and \LaTeX\2e. Their use is discouraged in expl3 (although still used in low-level definitions) because they are more fragile and in many cases require more expansion control (hence more code) than the two types of conditionals described above.

\begin{align*}
\texttt{\textbackslash c\_true\_bool} & & \text{Constants that represent } \texttt{true} \text{ and } \texttt{false}, \text{ respectively. Used to implement predicates.} \\
\texttt{\textbackslash c\_false\_bool}
\end{align*}

## 6.1 Tests on control sequences

\begin{align*}
\texttt{\textbackslash cs\_if\_eq\_p:NN} & & \texttt{\textbackslash cs\_if\_eq\_p:NN \langle cs \_1 \rangle \langle cs \_2 \rangle} \\
\texttt{\textbackslash cs\_if\_eq\_NNTF} & & \texttt{\textbackslash cs\_if\_eq\_NNTF \langle cs \_1 \rangle \langle cs \_2 \rangle \{\text{true code}\} \{\text{false code}\}}
\end{align*}

Compares the definition of two \texttt{(control sequences)} and is logically \texttt{true} if they are the same, \textit{i.e.} if they have exactly the same definition when examined with \texttt{\cs\_show\_N}.

\begin{align*}
\texttt{\textbackslash cs\_if\_exist\_p:N} & & \texttt{\textbackslash cs\_if\_exist\_p:N \langle control\ sequence \rangle} \\
\texttt{\textbackslash cs\_if\_exist\_p:c} & & \texttt{\textbackslash cs\_if\_exist\_p:c \langle control\ sequence \rangle \{\text{true code}\} \{\text{false code}\}} \\
\texttt{\textbackslash cs\_if\_exist\_NTF} & & \texttt{\textbackslash cs\_if\_exist\_NTF \langle control\ sequence \rangle \{\text{true code}\} \{\text{false code}\}} \\
\texttt{\textbackslash cs\_if\_exist\_cTF} & & \texttt{\textbackslash cs\_if\_exist\_cTF \langle control\ sequence \rangle \{\text{true code}\} \{\text{false code}\}}
\end{align*}

Tests whether the \texttt{(control sequence)} is currently defined (whether as a function or another control sequence type). Any definition of \texttt{(control sequence)} other than \texttt{\relax} evaluates as \texttt{true}.

\begin{align*}
\texttt{\textbackslash cs\_if\_free\_p:N} & & \texttt{\textbackslash cs\_if\_free\_p:N \langle control\ sequence \rangle} \\
\texttt{\textbackslash cs\_if\_free\_p:c} & & \texttt{\textbackslash cs\_if\_free\_p:c \langle control\ sequence \rangle \{\text{true code}\} \{\text{false code}\}} \\
\texttt{\textbackslash cs\_if\_free\_NTF} & & \texttt{\textbackslash cs\_if\_free\_NTF \langle control\ sequence \rangle \{\text{true code}\} \{\text{false code}\}} \\
\texttt{\textbackslash cs\_if\_free\_cTF} & & \texttt{\textbackslash cs\_if\_free\_cTF \langle control\ sequence \rangle \{\text{true code}\} \{\text{false code}\}}
\end{align*}

Tests whether the \texttt{(control sequence)} is currently free to be defined. This test is \texttt{false} if the \texttt{(control sequence)} currently exists (as defined by \texttt{\cs\_if\_exist\_N}).

## 6.2 Primitive conditionals

The \TeX engine itself provides many different conditionals. Some expand whatever comes after them and others don’t. Hence the names for these underlying functions often contains a \texttt{:w} part but higher level functions are often available. See for instance \texttt{\int\_compare\_p:nNn} which is a wrapper for \texttt{\if\_int\_compare\_w}.

Certain conditionals deal with specific data types like boxes and fonts and are described there. The ones described below are either the universal conditionals or deal with control sequences. We prefix primitive conditionals with \texttt{\if\_}.
\if_true: * \if_true: {true code} \else: {false code} \fi:
\if_false: * \if_false: {true code} \else: {false code} \fi:
\else: * \reverse_if:N {primitive conditional}
\fi: * \if_true: always executes {true code}, while \if_false: always executes {false code}.
\reverse_if:N reverses any two-way primitive conditional. \else: and \fi: delimit
the branches of the conditional. The function \or: is documented in \Bint and used in
case switches.

\textbf{\LaTeX hackers note:} These are equivalent to their corresponding \TeX primitive conditionals;
\reverse_if:N is \TeX’s \unless.

\if_meaning:w * \if_meaning:w \langle arg_1 \rangle \langle arg_2 \rangle {true code} \else: {false code} \fi:
\if:w * \if:w \langle token_1 \rangle \langle token_2 \rangle {true code} \else: {false code} \fi:
\if_charcode:w * \if_charcode:w \langle token_1 \rangle \langle token_2 \rangle {true code} \else: {false code} \fi:
\if_catcode:w * \if_catcode:w \langle token_1 \rangle \langle token_2 \rangle {true code} \else: {false code} \fi:

These conditionals expand any following tokens until two unexpandable tokens are left. If you wish to prevent this expansion, prefix the token in question with \exp_not:N. \if_catcode:w tests if the category codes of the two tokens are the same whereas \if:w tests if the character codes are identical. \if_charcode:w is an alternative name for \if:w.

\if_cs_exist:N * \if_cs_exist:N \langle cs \rangle {true code} \else: {false code} \fi:
\if_cs_exist:w * \if_cs_exist:w \langle tokens \rangle \cs_end: {true code} \else: {false code} \fi:

Check if \langle cs \rangle appears in the hash table or if the control sequence that can be formed from \langle tokens \rangle appears in the hash table. The latter function does not turn the control sequence in question into \scan_stop:! This can be useful when dealing with control sequences which cannot be entered as a single token.

\if_mode_horizontal: * \if_mode_horizontal: {true code} \else: {false code} \fi:
\if_mode_vertical: * \if_mode_vertical: {true code} \else: {false code} \fi:
\if_mode_math: * \if_mode_math: {true code} \else: {false code} \fi:
\if_mode_inner: * \if_mode_inner: {true code} \else: {false code} \fi:

Execute \langle true code \rangle if currently in horizontal mode, otherwise execute \langle false code \rangle. Similar for the other functions.
7 Starting a paragraph

\mode_leave_vertical:

Ensures that \TeX{} is not in vertical (inter-paragraph) mode. In horizontal or math mode this command has no effect, in vertical mode it switches to horizontal mode, and inserts a box of width \parindent, followed by the \everypar token list.

\TeX{}hackers note: This results in the contents of the \everypar token register being inserted, after \mode_leave_vertical: is complete. Notice that in contrast to the \LaTeX{}2ε leavevmode approach, no box is used by the method implemented here.

7.1 Debugging support

\debug_on:n \debug_off:n

\debug_on:n \{ \text{comma-separated list} \} \debug_off:n \{ \text{comma-separated list} \}

Turn on and off within a group various debugging code, some of which is also available as expl3 load-time options. The items that can be used in the \{\text{list}\} are

- \text{check-declarations} that checks all expl3 variables used were previously declared and that local/global variables (based on their name or on their first assignment) are only locally/globally assigned;

- \text{check-expressions} that checks integer, dimension, skip, and muskip expressions are not terminated prematurely;

- \text{deprecation} that makes soon-to-be-deprecated commands produce errors;

- \text{log-functions} that logs function definitions;

- \text{all} that does all of the above.

Providing these as switches rather than options allows testing code even if it relies on other packages: load all other packages, call \debug_on:n, and load the code that one is interested in testing. These functions can only be used in \LaTeX{}2ε package mode loaded with \enabledebug or another option implying it.

\debug_suspend: \debug_resume:

\debug_suspend: \ldots \debug_resume:

Suppress (locally) errors and logging from debug commands, except for the deprecation errors or warnings. These pairs of commands can be nested. This can be used around pieces of code that are known to fail checks, if such failures should be ignored. See for instance 13coffins.
Part V
The l3expansion package
Argument expansion

This module provides generic methods for expanding \TeX arguments in a systematic manner. The functions in this module all have prefix `exp`.

Not all possible variations are implemented for every base function. Instead only those that are used within the \LaTeX3 kernel or otherwise seem to be of general interest are implemented. Consult the module description to find out which functions are actually defined. The next section explains how to define missing variants.

1 Defining new variants

The definition of variant forms for base functions may be necessary when writing new functions or when applying a kernel function in a situation that we haven’t thought of before.

Internally preprocessing of arguments is done with functions of the form \verb|\exp_...| They all look alike, an example would be \verb|\exp_args:NNo|. This function has three arguments, the first and the second are a single tokens, while the third argument should be given in braces. Applying \verb|\exp_args:NNo| expands the content of third argument once before any expansion of the first and second arguments. If \verb|\seq_gpush:No| was not defined it could be coded in the following way:

\begin{verbatim}
\exp_args:NNo \seq_gpush:Nn
\g_file_name_stack
{ \l_tmpa_tl }
\end{verbatim}

In other words, the first argument to \verb|\exp_args:NNo| is the base function and the other arguments are preprocessed and then passed to this base function. In the example the first argument to the base function should be a single token which is left unchanged while the second argument is expanded once. From this example we can also see how the variants are defined. They just expand into the appropriate \verb|\exp_| function followed by the desired base function, e.g.

\begin{verbatim}
\cs_generate_variant:Nn \seq_gpush:Nn { No }
\end{verbatim}

results in the definition of \verb|\seq_gpush:No|

\begin{verbatim}
\cs_new:Npn \seq_gpush:No { \exp_args:NNo \seq_gpush:Nn }
\end{verbatim}

Providing variants in this way in style files is safe as the \verb|\cs_generate_variant:Nn| function will only create new definitions if there is not already one available. Therefore adding such definition to later releases of the kernel will not make such style files obsolete.

The steps above may be automated by using the function \verb|\cs_generate_variant:Nn|, described next.
2 Methods for defining variants

We recall the set of available argument specifiers.

- \( N \) is used for single-token arguments while \( c \) constructs a control sequence from its name and passes it to a parent function as an \( N \)-type argument.

- Many argument types extract or expand some tokens and provide it as an \( n \)-type argument, namely a braced multiple-token argument: \( V \) extracts the value of a variable, \( v \) extracts the value from the name of a variable, \( n \) uses the argument as it is, \( o \) expands once, \( f \) expands fully the front of the token list, \( e \) and \( x \) expand fully all tokens (differences are explained later).

- A few odd argument types remain: \( T \) and \( F \) for conditional processing, otherwise identical to \( n \)-type arguments, \( p \) for the parameter text in definitions, \( w \) for arguments with a specific syntax, and \( D \) to denote primitives that should not be used directly.
This function is used to define argument-specifier variants of the \texttt{\emph{parent control sequence}} for \TeX{} code-level macros. The \texttt{\emph{parent control sequence}} is first separated into the \texttt{\emph{base name}} and \texttt{\emph{original argument specifier}}. The comma-separated list of \texttt{\emph{variant argument specifiers}} is then used to define variants of the \texttt{\emph{original argument specifier}} if these are not already defined. For each \texttt{\emph{variant}} given, a function is created that expands its arguments as detailed and passes them to the \texttt{\emph{parent control sequence}}. So for example

\begin{verbatim}
  \cs_set:Npn \foo:Nn #1#2 { code here }
  \cs_generate_variant:Nn \foo:Nn { c }
\end{verbatim}

creates a new function \texttt{\foo:cn} which expands its first argument into a control sequence name and passes the result to \texttt{\foo:Nn}. Similarly

\begin{verbatim}
  \cs_set:Npn \foo:Nn #1#2 { code here }
  \cs_generate_variant:Nn \foo:Nn { NV , cV }
\end{verbatim}

generates the functions \texttt{\foo:NV} and \texttt{\foo:cV} in the same way. The \texttt{\cs_generate_variant:Nn} function can only be applied if the \texttt{\emph{parent control sequence}} is already defined. If the \texttt{\emph{parent control sequence}} is protected or if the \texttt{\emph{variant}} involves any \texttt{x} argument, then the \texttt{\emph{variant control sequence}} is also protected. The \texttt{\emph{variant}} is created globally, as is any \texttt{\exp_args:No} function needed to carry out the expansion.

Only \texttt{n} and \texttt{N} arguments can be changed to other types. The only allowed changes are

- \texttt{c} variant of an \texttt{N} parent;
- \texttt{o, V, v, f, e, or x} variant of an \texttt{n} parent;
- \texttt{N, n, T, F, or p} argument unchanged.

This means the \texttt{\emph{parent}} of a \texttt{\emph{variant}} form is always unambiguous, even in cases where both an \texttt{n}-type parent and an \texttt{N}-type parent exist, such as for \texttt{\tl_count:n} and \texttt{\tl_count:N}.

For backward compatibility it is currently possible to make \texttt{n, o, V, v, f, e, or x}-type variants of an \texttt{N}-type argument or \texttt{N} or \texttt{c}-type variants of an \texttt{n}-type argument. Both are deprecated. The first because passing more than one token to an \texttt{N}-type argument will typically break the parent function’s code. The second because programmers who use that most often want to access the value of a variable given its name, hence should use a \texttt{V}-type or \texttt{v}-type variant instead of \texttt{c}-type. In those cases, using the lower-level \texttt{\exp_args:No} or \texttt{\exp_args:Nc} functions explicitly is preferred to defining confusing variants.

### 3 Introducing the variants

The \texttt{V} type returns the value of a register, which can be one of \texttt{tl, clist, int, skip, dim, muskip}, or built-in \TeX{} registers. The \texttt{v} type is the same except it first creates a control sequence out of its argument before returning the value.

In general, the programmer should not need to be concerned with expansion control. When simply using the content of a variable, functions with a \texttt{V} specifier should be used. For those referred to by (cs)name, the \texttt{v} specifier is available for the same purpose. Only
when specific expansion steps are needed, such as when using delimited arguments, should the lower-level functions with \o specifiers be employed.

The \e type expands all tokens fully, starting from the first. More precisely the expansion is identical to that of \TeX’s \textbackslash message (in particular \# needs not be doubled). It was added in May 2018. In recent enough engines (starting around 2019) it relies on the primitive \expanded hence is fast. In older engines it is very much slower. As a result it should only be used in performance critical code if typical users will have a recent installation of the \TeX ecosystem.

The \x type expands all tokens fully, starting from the first. In contrast to \e, all macro parameter characters \# must be doubled, and omitting this leads to low-level errors. In addition this type of expansion is not expandable, namely functions that have \x in their signature do not themselves expand when appearing inside \x or \e expansion.

The \f type is so special that it deserves an example. It is typically used in contexts where only expandable commands are allowed. Then \x-expansion cannot be used, and \f-expansion provides an alternative that expands the front of the token list as much as can be done in such contexts. For instance, say that we want to evaluate the integer expression \(3 + 4\) and pass the result 7 as an argument to an expandable function \texttt{\example:n}. For this, one should define a variant using \texttt{\cs_generate_variant:Nn \example:n { f }}, then do

\begin{verbatim}
\example:f { \int_eval:n { 3 + 4 } }
\end{verbatim}

Note that \x-expansion would also expand \texttt{\int_eval:n} fully to its result 7, but the variant \texttt{\example:x} cannot be expandable. Note also that \o-expansion would not expand \texttt{\int_eval:n} fully to its result since that function requires several expansions. Besides the fact that \x-expansion is protected rather than expandable, another difference between \f-expansion and \x-expansion is that \f-expansion expands tokens from the beginning and stops as soon as a non-expandable token is encountered, while \x-expansion continues expanding further tokens. Thus, for instance

\begin{verbatim}
\example:f { \int_eval:n { 1 + 2 } , \int_eval:n { 3 + 4 } }
\end{verbatim}

results in the call

\begin{verbatim}
\example:n { 3 , \int_eval:n { 3 + 4 } }
\end{verbatim}

while using \texttt{\example:x} or \texttt{\example:e} instead results in

\begin{verbatim}
\example:n { 3 , 7 }
\end{verbatim}

at the cost of being protected (for \x type) or very much slower in old engines (for \e type). If you use \f type expansion in conditional processing then you should stick to using \TF type functions only as the expansion does not finish any \texttt{\if... \fi}: itself!

It is important to note that both \f- and \o-type expansion are concerned with the expansion of tokens from left to right in their arguments. In particular, \o-type expansion applies to the first \texttt{token} in the argument it receives: it is conceptually similar to

\begin{verbatim}
\exp_after:wN <base function> \exp_after:wN { <argument> }
\end{verbatim}

At the same time, \f-type expansion stops at the first non-expandable token. This means for example that both

\begin{verbatim}
\tl_set:No \l_tmpa_tl { { \g_tmpb_tl } }
\end{verbatim}

28
and
\tl_set:Nf \l_tmpa_tl { { \g_tmpb_tl } }
leave \g_tmpb_tl unchanged: { is the first token in the argument and is non-expandable.
It is usually best to keep the following in mind when using variant forms.

- Variants with x-type arguments (that are fully expanded before being passed to
  the n-type base function) are never expandable even when the base function is.
  Such variants cannot work correctly in arguments that are themselves subject to
  expansion. Consider using f or e expansion.
- In contrast, e expansion (full expansion, almost like x except for the treatment of #)
  does not prevent variants from being expandable (if the base function is). The draw-
  back is that e expansion is very much slower in old engines (before 2019). Consider
  using f expansion if that type of expansion is sufficient to perform the required
  expansion, or x expansion if the variant will not itself need to be expandable.
- Finally f expansion only expands the front of the token list, stopping at the first
  non-expandable token. This may fail to fully expand the argument.

When speed is essential (for functions that do very little work and whose variants are
used numerous times in a document) the following considerations apply because internal
functions for argument expansion come in two flavours, some faster than others.

- Arguments that might need expansion should come first in the list of arguments.
- Arguments that should consist of single tokens \n, \c, \V, or \v should come first among
  these.
- Arguments that appear after the first multi-token argument \n, \f, \e, or \o require
  slightly slower special processing to be expanded. Therefore it is best to use the
  optimized functions, namely those that contain only \n, \c, \V, and \v, and, in the last
  position, \o, \f, \e, with possible trailing \N or \n or \T or \F, which are not expanded.
  Any x-type argument causes slightly slower processing.

## 4 Manipulating the first argument

These functions are described in detail: expansion of multiple tokens follows the same
rules but is described in a shorter fashion.

\exp_args:Nc \exp_args:Nc \exp_args:cc

This function absorbs two arguments (the (function) name and the (tokens)). The
(tokens) are expanded until only characters remain, and are then turned into a control
sequence. The result is inserted into the input stream after reinsertion of the (function).
Thus the (function) may take more than one argument: all others are left unchanged.

The :cc variant constructs the (function) name in the same manner as described for
the (tokens).

**\TeXhackers note:** Protected macros that appear in a c-type argument are expanded
despite being protected; \exp_not:n also has no effect. An internal error occurs if non-characters
or active characters remain after full expansion, as the conversion to a control sequence is not
possible.
This function absorbs two arguments (the \( \text{function} \) name and the \( \text{tokens} \)). The \( \text{tokens} \) are expanded once, and the result is inserted in braces into the input stream \emph{after} reinsertion of the \( \text{function} \). Thus the \( \text{function} \) may take more than one argument: all others are left unchanged.

This function absorbs two arguments (the names of the \( \text{function} \) and the \( \text{variable} \)). The content of the \( \text{variable} \) are recovered and placed inside braces into the input stream \emph{after} reinsertion of the \( \text{function} \). Thus the \( \text{function} \) may take more than one argument: all others are left unchanged.

This function absorbs two arguments (the \( \text{function} \) name and the \( \text{tokens} \)). The \( \text{tokens} \) are expanded until only characters remain, and are then turned into a control sequence. This control sequence should be the name of a \( \text{variable} \). The content of the \( \text{variable} \) are recovered and placed inside braces into the input stream \emph{after} reinsertion of the \( \text{function} \). Thus the \( \text{function} \) may take more than one argument: all others are left unchanged.

\textbf{\TeXhacker{s} note:} Protected macros that appear in a \texttt{v}-type argument are expanded despite being protected; \texttt{\exp_not:n} also has no effect. An internal error occurs if non-characters or active characters remain after full expansion, as the conversion to a control sequence is not possible.

This function absorbs two arguments (the \( \text{function} \) name and the \( \text{tokens} \)) and exhaustively expands the \( \text{tokens} \). The result is inserted in braces into the input stream \emph{after} reinsertion of the \( \text{function} \). Thus the \( \text{function} \) may take more than one argument: all others are left unchanged.

\textbf{\TeXhacker{s} note:} This relies on the \texttt{\expanded} primitive when available (in \LaTeX{} and starting around 2019 in other engines). Otherwise it uses some fall-back code that is very much slower. As a result it should only be used in performance-critical code if typical users have a recent installation of the \TeX{} ecosystem.

This function absorbs two arguments (the \( \text{function} \) name and the \( \text{tokens} \)). The \( \text{tokens} \) are fully expanded until the first non-expandable token is found (if that is a space it is removed), and the result is inserted in braces into the input stream \emph{after} reinsertion of the \( \text{function} \). Thus the \( \text{function} \) may take more than one argument: all others are left unchanged.
This function absorbs two arguments (the `function` name and the `tokens`) and exhaustively expands the `tokens`. The result is inserted in braces into the input stream after reinsertion of the `function`. Thus the `function` may take more than one argument: all others are left unchanged.

5 Manipulating two arguments

These optimized functions absorb three arguments and expand the second and third as detailed by their argument specifier. The first argument of the function is then the next item on the input stream, followed by the expansion of the second and third arguments.

These functions absorb three arguments and expand the second and third as detailed by their argument specifier. The first argument of the function is then the next item on the input stream, followed by the expansion of the second and third arguments. These functions need slower processing.

These functions absorb three arguments and expand the second and third as detailed by their argument specifier. The first argument of the function is then the next item on the input stream, followed by the expansion of the second and third arguments. These functions are not expandable due to their x-type argument.
6 Manipulating three arguments

\texttt{\exp_args:NNNo} \* \\
\texttt{\exp_args:NNNV} \* \\
\texttt{\exp_args:Nccc} \* \\
\texttt{\exp_args:NcNc} \* \\
\texttt{\exp_args:NcNo} \* \\
\texttt{\exp_args:Ncco} "

These optimized functions absorb four arguments and expand the second, third and fourth as detailed by their argument specifier. The first argument of the function is then the next item on the input stream, followed by the expansion of the second argument, \textit{etc.}

\texttt{\exp_args:NNcf} \* \\
\texttt{\exp_args:NNno} \* \\
\texttt{\exp_args:NNnV} \* \\
\texttt{\exp_args:NNVV} \* \\
\texttt{\exp_args:Ncno} \* \\
\texttt{\exp_args:NcnV} \* \\
\texttt{\exp_args:Ncoo} \* \\
\texttt{\exp_args:NcVV} \* \\
\texttt{\exp_args:Nmnc} \* \\
\texttt{\exp_args:Nmno} \* \\
\texttt{\exp_args:Nmnf} \* \\
\texttt{\exp_args:Naff} \* \\
\texttt{\exp_args:Nooo} \* \\
\texttt{\exp_args:Nooof} \* \\
\texttt{\exp_args:Nffo} \* \\
\texttt{\exp_args:Neee} "

\texttt{\exp_args:NNx} \* \\
\texttt{\exp_args:NNnx} \* \\
\texttt{\exp_args:NNox} \* \\
\texttt{\exp_args:Nccx} \* \\
\texttt{\exp_args:NcNx} \* \\
\texttt{\exp_args:Nmnx} \* \\
\texttt{\exp_args:Nmox} \* \\
\texttt{\exp_args:Noox} \* \\

\texttt{\exp_args:NNx} \texttt{(token_1 \ (token_2 \ (token_3} \ \{\tokens\})}

These functions absorb four arguments and expand the second, third and fourth as detailed by their argument specifier. The first argument of the function is then the next item on the input stream, followed by the expansion of the second argument, \textit{etc.}

\texttt{\exp_args:NNx} \texttt{(token_1 \ (token_2 \ (tokens_1} \ \{\tokens_2\})}

These functions absorb four arguments and expand the second, third and fourth as detailed by their argument specifier. The first argument of the function is then the next item on the input stream, followed by the expansion of the second argument, \textit{etc.}

New: 2015-08-12
7 Unbraced expansion

\exp_last_unbraced:Nno \exp_last_unbraced:Noo \exp_last_unbraced:Nfo \exp_last_unbraced:NNNo
\exp_last_unbraced:Nx

These functions absorb the number of arguments given by their specification, carry out the expansion indicated and leave the results in the input stream, with the last argument not surrounded by the usual braces. Of these, the :Nno, :Noo, :Nfo and :NnNo variants need slower processing.

\TeXhacks\note: As an optimization, the last argument is unbraced by some of those functions before expansion. This can cause problems if the argument is empty: for instance, \code{\exp_last_unbraced:Nf \foo_bar:w { \q_stop} leads to an infinite loop, as the quark is \texttt{f}-expanded.

\exp_last_unbraced:Nv \exp_last_unbraced:Nf \exp_last_unbraced:NNo \exp_last_unbraced:NNV \exp_last_unbraced:Nff \exp_last_unbraced:NfNf

\exp_last_unbraced:No \exp_last_unbraced:NV \exp_last_unbraced:Nf \exp_last_unbraced:NNo \exp_last_unbraced:NNf \exp_last_unbraced:NNf

\exp_last_unbraced:Nco \exp_last_unbraced:NcV \exp_last_unbraced:Nno \exp_last_unbraced:Noo \exp_last_unbraced:Nfo \exp_last_unbraced:NfNf

\exp_last_unbraced:NNo \exp_last_unbraced:NNf \exp_last_unbraced:NnNo \exp_last_unbraced:NNNNf

\exp_last_unbraced:NNNo \exp_last_unbraced:NNNV \exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

\exp_last_unbraced:NNf

Updated: 2018-05-15

\exp_last_unbraced:Nx \langle \text{function} \rangle \{ \langle \text{tokens} \rangle \}

This function fully expands the \langle \text{tokens} \rangle and leaves the result in the input stream after reinsertion of the \langle \text{function} \rangle. This function is not expandable.

\exp_last_two_unbraced:Noo \exp_last_two_unbraced:Noo \langle \text{token} \rangle \{ \langle \text{tokens} \rangle \} \{ \langle \text{tokens} \rangle \}

This function absorbs three arguments and expands the second and third once. The first argument of the function is then the next item on the input stream, followed by the expansion of the second and third arguments, which are not wrapped in braces. This function needs special (slower) processing.

\exp_after:wN \langle \text{token} \rangle \langle \text{token} \rangle

Carries out a single expansion of \langle \text{token} \rangle (which may consume arguments) prior to the expansion of \langle \text{token} \rangle. If \langle \text{token} \rangle has no expansion (for example, if it is a character) then it is left unchanged. It is important to notice that \langle \text{token} \rangle may be any single token, including group-opening and -closing tokens \{ or \} assuming normal \TeX\ category codes. Unless specifically required this should be avoided: expansion should be carried out using an appropriate argument specifier variant or the appropriate \texttt{\exp_arg:N} function.

\TeXhacks\note: This is the \TeX\ primitive \texttt{\expandafter} renamed.

8 Preventing expansion

Despite the fact that the following functions are all about preventing expansion, they’re designed to be used in an expandable context and hence are all marked as being \texttt{expandable}. 
able’ since they themselves disappear after the expansion has completed.

\exp_not:N \( \langle \text{token} \rangle \)

Prevents expansion of the \( \langle \text{token} \rangle \) in a context where it would otherwise be expanded, for example an \( x \)-type argument or the first token in an \( o \) or \( e \) or \( f \) argument.

\textbf{TEXhackers note:} This is the \TeX\ \texttt{noexpand} primitive. It only prevents expansion. At the beginning of an \( f \)-type argument, a space \( \langle \text{token} \rangle \) is removed even if it appears as \exp_not:N \c_space_token. In an \( x \)-expanding definition (\cs_new:Npx), a macro parameter introduces an argument even if it appears as \exp_not:N \# 1. This differs from \exp_not:n.

\exp_not:c \( \langle \text{tokens} \rangle \)

Expands the \( \langle \text{tokens} \rangle \) until only characters remain, and then converts this into a control sequence. Further expansion of this control sequence is then inhibited using \exp_not:N.

\textbf{TEXhackers note:} Protected macros that appear in a \( c \)-type argument are expanded despite being protected; \exp_not:n also has no effect. An internal error occurs if non-characters or active characters remain after full expansion, as the conversion to a control sequence is not possible.

\exp_not:n \( \langle \text{tokens} \rangle \)

Prevents expansion of the \( \langle \text{tokens} \rangle \) in an \( e \) or \( x \)-type argument. In all other cases the \( \langle \text{tokens} \rangle \) continue to be expanded, for example in the input stream or in other types of arguments such as \( c, f, v \). The argument of \exp_not:n must be surrounded by braces.

\textbf{TEXhackers note:} This is the \( \varepsilon \)-\TeX\ \texttt{unexpanded} primitive. In an \( x \)-expanding definition (\cs_new:Npx), \exp_not:n \( \#1 \) is equivalent to \#\#1 rather than to \#1, namely it inserts the two characters \# and \#1. In an \( e \)-type argument \exp_not:n \( \# \) is equivalent to \#, namely it inserts the character \#.

\exp_not:o \( \langle \text{tokens} \rangle \)

Expands the \( \langle \text{tokens} \rangle \) once, then prevents any further expansion in \( x \)-type or \( e \)-type arguments using \exp_not:n.

\exp_not:V \( \langle \text{variable} \rangle \)

Recovers the content of the \( \langle \text{variable} \rangle \), then prevents expansion of this material in \( x \)-type or \( e \)-type arguments using \exp_not:n.
\exp_not:v \{\langle tokens\rangle\}

Expands the \langle tokens\rangle until only characters remains, and then converts this into a control sequence which should be a \langle variable\rangle name. The content of the \langle variable\rangle is recovered, and further expansion in x-type or e-type arguments is prevented using \exp_not:n.

\textbf{\LaTeX} hackers note: Protected macros that appear in a v-type argument are expanded despite being protected; \exp_not:n also has no effect. An internal error occurs if non-characters or active characters remain after full expansion, as the conversion to a control sequence is not possible.

\exp_not:e \{\langle tokens\rangle\}

Expands \langle tokens\rangle exhaustively, then protects the result of the expansion (including any tokens which were not expanded) from further expansion in e or x-type arguments using \exp_not:n. This is very rarely useful but is provided for consistency.

\exp_not:f \{\langle tokens\rangle\}

Expands \langle tokens\rangle fully until the first unexpandable token is found (if it is a space it is removed). Expansion then stops, and the result of the expansion (including any tokens which were not expanded) is protected from further expansion in x-type or e-type arguments using \exp_not:n.

\exp_stop_f: \{\langle tokens\rangle \exp_stop_f: \langle more tokens\rangle\}

This function terminates an f-type expansion. Thus if a function \foo_bar:f starts an f-type expansion and all of \langle tokens\rangle are expandable \exp_stop_f: terminates the expansion of tokens even if \langle more tokens\rangle are also expandable. The function itself is an implicit space token. Inside an x-type expansion, it retains its form, but when typeset it produces the underlying space (\uni{2423}).

9 Controlled expansion

The \exp^3 language makes all efforts to hide the complexity of \LaTeX expansion from the programmer by providing concepts that evaluate/expand arguments of functions prior to calling the “base” functions. Thus, instead of using many \expandafter calls and other trickery it is usually a matter of choosing the right variant of a function to achieve a desired result.

Of course, deep down \LaTeX is using expansion as always and there are cases where a programmer needs to control that expansion directly; typical situations are basic data manipulation tools. This section documents the functions for that level. These commands are used throughout the kernel code, but we hope that outside the kernel there will be little need to resort to them. Instead the argument manipulation methods document above should usually be sufficient.

While \exp_after:wN expands one token (out of order) it is sometimes necessary to expand several tokens in one go. The next set of commands provide this functionality. Be aware that it is absolutely required that the programmer has full control over the tokens to be expanded, i.e., it is not possible to use these functions to expand unknown input as part of \langle expandable-tokens\rangle as that will break badly if unexpandable tokens are encountered in that place!
\exp:w \exp:w \exp:w \exp:w \exp_end:: *

\exp:w \langle \text{expandable tokens} \rangle \exp_end:

Expands \langle \text{expandable-tokens} \rangle until reaching \exp_end: at which point expansion stops. The full expansion of \langle \text{expandable tokens} \rangle has to be empty. If any token in \langle \text{expandable tokens} \rangle or any token generated by expanding the tokens therein is not expandable the expansion will end prematurely and as a result \exp_end: will be misinterpreted later on.\footnote{Due to the implementation you might get the character in position 0 in the current font (typically ‘"’) in the output without any error message!}

In typical use cases the \exp_end: is hidden somewhere in the replacement text of \langle \text{expandable-tokens} \rangle rather than being on the same expansion level than \exp:w, e.g., you may see code such as

\exp:w \texttt{\@@_case:NnTF} \#1 \{ \} \{ \}

where somewhere during the expansion of \texttt{\@@_case:NnTF} the \exp_end: gets generated.

\TeXhackers note: The current implementation uses \texttt{\romannumeral} hence ignores space tokens and explicit signs + and - in the expansion of the \langle \text{expandable tokens} \rangle, but this should not be relied upon.

\begin{verbatim}
\exp:w \exp:w \exp_end:: *
\exp:w \langle \text{expandable-tokens} \rangle \exp_end_continue_f:w \langle \text{further-tokens} \rangle
\end{verbatim}

Expands \langle \text{expandable-tokens} \rangle until reaching \exp_end_continue_f:w at which point expansion continues as an \texttt{f}-type expansion expanding \langle \text{further-tokens} \rangle until an unexpandable token is encountered (or the \texttt{f}-type expansion is explicitly terminated by \exp_stop_f:). As with all \texttt{f}-type expansions a space ending the expansion gets removed.

The full expansion of \langle \text{expandable-tokens} \rangle has to be empty. If any token in \langle \text{expandable-tokens} \rangle or any token generated by expanding the tokens therein is not expandable the expansion will end prematurely and as a result \exp_end_continue_f:w will be misinterpreted later on.\footnote{In this particular case you may get a character into the output as well as an error message.}

In typical use cases \langle \text{expandable-tokens} \rangle contains no tokens at all, e.g., you will see code such as

\exp_after:wN \{ \exp:w \exp_end_continue_f:w \langle \text{further-tokens} \rangle \exp_stop_f: \#2 \}

where the \exp_after:wN triggers an \texttt{f}-expansion of the tokens in \#2. For technical reasons this has to happen using two tokens (if they would be hidden inside another command \exp_after:wN would only expand the command but not trigger any additional \texttt{f}-expansion).

You might wonder why there are two different approaches available, after all the effect of

\exp:w \langle \text{expandable-tokens} \rangle \exp_end:

\exp:w \exp_end_continue_f:w \langle \text{expandable-tokens} \rangle \exp_stop_f:

The reason is simply that the first approach is slightly faster (one less token to parse and less expansion internally) so in places where such performance really matters and where we want to explicitly stop the expansion at a defined point the first form is preferable.
The difference to \exp_end_continue_f:w is that we first we pick up an argument which is then returned to the input stream. If \langle further-tokens \rangle starts with space tokens then these space tokens are removed while searching for the argument. If it starts with a brace group then the braces are removed. Thus such spaces or braces will not terminate the f-type expansion.

10 Internal functions

\cs_new:Npn \exp_args:Ncof { \::c \::o \::f \::: }
\cs_new:Npn \exp_last_unbraced:Nno { \::n \::o_unbraced \::: }

Internal forms for the base expansion types. These names do not conform to the general \LaTeX3 approach as this makes them more readily visible in the log and so forth. They should not be used outside this module.

\cs_new:Npn \exp: \langle expandable-tokens \rangle \exp_end_continue_f:nw \langle further-tokens \rangle
\cs_new:Npn \exp_last_unbraced:Nno { \::n \::o_unbraced \::: }

Internal forms for the expansion types which leave the terminal argument unbraced. These names do not conform to the general \LaTeX3 approach as this makes them more readily visible in the log and so forth. They should not be used outside this module.
Part VI
The l3tl package
Token lists

\TeX\ works with tokens, and \LaTeX3\ therefore provides a number of functions to deal with lists of tokens. Token lists may be present directly in the argument to a function:

\begin{verbatim}
\foo:n { a collection of \tokens }
\end{verbatim}

or may be stored in a so-called “token list variable”, which have the suffix tl: a token list variable can also be used as the argument to a function, for example

\begin{verbatim}
\foo:N \l_{\text{some\_tl}}
\end{verbatim}

In both cases, functions are available to test and manipulate the lists of tokens, and these have the module prefix tl. In many cases, functions which can be applied to token list variables are paired with similar functions for application to explicit lists of tokens: the two “views” of a token list are therefore collected together here.

A token list (explicit, or stored in a variable) can be seen either as a list of “items”, or a list of “tokens”. An item is whatever \use:n\ would grab as its argument: a single non-space token or a brace group, with optional leading explicit space characters (each item is thus itself a token list). A token is either a normal N argument, or \{, or \} (assuming normal \TeX\ category codes). Thus for example

\begin{verbatim}
{ Hello } \text{- world}
\end{verbatim}

contains six items (Hello, w, o, r, l and d), but thirteen tokens ({, H, e, l, o, }, \{, w, o, r, l and d). Functions which act on items are often faster than their analogue acting directly on tokens.

1 Creating and initialising token list variables

\begin{verbatim}
\tl_new:N \tl_new:c
\tl_const:Nn \tl_const:c(Nx|cn|cx)
\tl_clear:N \tl_clear:c \tl_gclear:N \tl_gclear:c
\end{verbatim}

\begin{verbatim}
\tl_new:N (tl var)
\tl_const:Nn (tl var) {\langle token list\rangle}
\tl_clear:N (tl var)
\end{verbatim}

\begin{verbatim}
Creates a new (tl var) or raises an error if the name is already taken. The declaration is global. The (tl var) is initially empty.

\tl_const:Nn (tl var) {\langle token list\rangle}

Creates a new constant (tl var) or raises an error if the name is already taken. The value of the (tl var) is set globally to the (token list).

\tl_clear:N (tl var)

Clears all entries from the (tl var).
\texttt{\textbackslash tl\_clear\_new:N} \texttt{\{tl var\}}
\texttt{\textbackslash tl\_clear\_new:c}
\texttt{\textbackslash tl\_gclear\_new:N}
\texttt{\textbackslash tl\_gclear\_new:c}

Ensures that the \langle tl var \rangle exists globally by applying \texttt{\textbackslash tl\_new:N} if necessary, then applies \texttt{\textbackslash tl\_(g)clear:N} to leave the \langle tl var \rangle empty.

\texttt{\textbackslash tl\_set\_eq:NN} \texttt{\langle tl var \rangle} \texttt{\langle tl var \rangle}

Sets the content of \langle tl var \rangle equal to that of \langle tl var \rangle.

\texttt{\textbackslash tl\_set\_eq:NN} \texttt{\langle tl var \rangle} \texttt{\langle tl var \rangle} \texttt{\langle tl var \rangle}

\texttt{\textbackslash tl\_gset\_eq:NN} \texttt{\langle tl var \rangle} \texttt{\langle tl var \rangle} \texttt{\langle tl var \rangle}

Concatenates the content of \langle tl var \rangle and \langle tl var \rangle to together and saves the result in \langle tl var \rangle. The \langle tl var \rangle is placed at the left side of the new token list.

\texttt{\textbackslash tl\_if\_exist:p:N} \texttt{\langle tl var \rangle} \texttt{\star}
\texttt{\textbackslash tl\_if\_exist:p:c} \texttt{\star}
\texttt{\textbackslash tl\_if\_exist:NFF} \texttt{\langle tl var \rangle} \texttt{\langle true code \rangle} \texttt{\langle false code \rangle} \texttt{\star}
\texttt{\textbackslash tl\_if\_exist:c} \texttt{\langle tl var \rangle} \texttt{\langle true code \rangle} \texttt{\langle false code \rangle} \texttt{\star}

Tests whether the \langle tl var \rangle is currently defined. This does not check that the \langle tl var \rangle really is a token list variable.

\section{Adding data to token list variables}

\texttt{\textbackslash tl\_set:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_set:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_gset:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_gset:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}

Sets \langle tl var \rangle to contain \langle tokens \rangle, removing any previous content from the variable.

\texttt{\textbackslash tl\_put\_left:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_put\_left:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_gput\_left:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_gput\_left:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}

Appends \langle tokens \rangle to the left side of the current content of \langle tl var \rangle.

\texttt{\textbackslash tl\_put\_right:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_put\_right:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_gput\_right:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}
\texttt{\textbackslash tl\_gput\_right:Nn} \texttt{\langle tl var \rangle} \texttt{\langle tokens \rangle}

Appends \langle tokens \rangle to the right side of the current content of \langle tl var \rangle.
3 Modifying token list variables

\texttt{\tl_replace_once:Nnn} \langle\texttt{tl var}\rangle \langle\texttt{old tokens}\rangle \langle\texttt{new tokens}\rangle

Replaces the first (leftmost) occurrence of \langle\texttt{old tokens}\rangle in the \langle\texttt{tl var}\rangle with \langle\texttt{new tokens}\rangle. \langle\texttt{Old tokens}\rangle cannot contain \{, \} or \# (more precisely, explicit character tokens with category code 1 (begin-group) or 2 (end-group), and tokens with category code 6).

\texttt{\tl_replace_all:Nnn} \langle\texttt{tl var}\rangle \langle\texttt{old tokens}\rangle \langle\texttt{new tokens}\rangle

Replaces all occurrences of \langle\texttt{old tokens}\rangle in the \langle\texttt{tl var}\rangle with \langle\texttt{new tokens}\rangle. \langle\texttt{Old tokens}\rangle cannot contain \{, \} or \# (more precisely, explicit character tokens with category code 1 (begin-group) or 2 (end-group), and tokens with category code 6). As this function operates from left to right, the pattern \langle\texttt{old tokens}\rangle may remain after the replacement (see \texttt{\tl_remove_all:Nn} for an example).

\texttt{\tl_remove_once:Nn} \langle\texttt{tl var}\rangle \langle\texttt{tokens}\rangle

Removes the first (leftmost) occurrence of \langle\texttt{tokens}\rangle from the \langle\texttt{tl var}\rangle. \langle\texttt{Tokens}\rangle cannot contain \{, \} or \# (more precisely, explicit character tokens with category code 1 (begin-group) or 2 (end-group), and tokens with category code 6).

\texttt{\tl_remove_all:Nn} \langle\texttt{tl var}\rangle \langle\texttt{tokens}\rangle

Removes all occurrences of \langle\texttt{tokens}\rangle from the \langle\texttt{tl var}\rangle. \langle\texttt{Tokens}\rangle cannot contain \{, \} or \# (more precisely, explicit character tokens with category code 1 (begin-group) or 2 (end-group), and tokens with category code 6). As this function operates from left to right, the pattern \langle\texttt{tokens}\rangle may remain after the removal, for instance,

\texttt{\tl_set:Nn \l_tmpa_tl {abbccd} \tl_remove_all:Nn \l_tmpa_tl \{bc\}}

results in \l_tmpa_tl containing abcd.

4 Reassigning token list category codes

These functions allow the rescanning of tokens: re-apply \TeX{}’s tokenization process to apply category codes different from those in force when the tokens were absorbed. Whilst this functionality is supported, it is often preferable to find alternative approaches to achieving outcomes rather than rescanning tokens (for example construction of token lists token-by-token with intervening category code changes or using \texttt{\char_generate:nn}).
\tl_set_rescan:Nnn \tl_set_rescan:(\NoNnx|\cnn|\cno|\cnx) 
\tl_gset_rescan:Nnn \tl_gset_rescan:(\NoNnx|\cnn|\cno|\cnx)

Sets \textit{\texttt{\langle tl var \rangle}} to contain \textit{\texttt{\langle tokens \rangle}}, applying the category code régime specified in the \textit{\texttt{\langle setup \rangle}} before carrying out the assignment. (Category codes applied to tokens not explicitly covered by the \textit{\texttt{\langle setup \rangle}} are those in force at the point of use of \texttt{\tl_set_rescan:Nnn}.) This allows the \textit{\texttt{\langle tl var \rangle}} to contain material with category codes other than those that apply when \textit{\texttt{\langle tokens \rangle}} are absorbed. The \textit{\texttt{\langle setup \rangle}} is run within a group and may contain any valid input, although only changes in category codes are relevant. See also \texttt{\tl_rescan:nn}.

\textbf{\texttt{\LaTeX}hackers note:} The \textit{\texttt{\langle tokens \rangle}} are first turned into a string (using \texttt{\tl_to_str:n}). If the string contains one or more characters with character code \texttt{\newlinechar} (set equal to \texttt{\endlinechar} unless that is equal to 32, before the user \textit{\texttt{\langle setup \rangle}}), then it is split into lines at these characters, then read as if reading multiple lines from a file, ignoring spaces (catcode 10) at the beginning and spaces and tabs (character code 32 or 9) at the end of every line. Otherwise, spaces (and tabs) are retained at both ends of the single-line string, as if it appeared in the middle of a line read from a file.

\texttt{\tl_rescan:nn} \texttt{(\langle setup \rangle)} \texttt{(\langle tokens \rangle)}

Rescans \textit{\texttt{\langle tokens \rangle}} applying the category code régime specified in the \textit{\texttt{\langle setup \rangle}}, and leaves the resulting tokens in the input stream. (Category codes applied to tokens not explicitly covered by the \textit{\texttt{\langle setup \rangle}} are those in force at the point of use of \texttt{\tl_rescan:nn}.) The \textit{\texttt{\langle setup \rangle}} is run within a group and may contain any valid input, although only changes in category codes are relevant. See also \texttt{\tl_set_rescan:Nnn}, which is more robust than using \texttt{\tl_set:Nn} in the \textit{\texttt{\langle tokens \rangle}} argument of \texttt{\tl_rescan:nn}.

\textbf{\texttt{\LaTeX}hackers note:} The \textit{\texttt{\langle tokens \rangle}} are first turned into a string (using \texttt{\tl_to_str:n}). If the string contains one or more characters with character code \texttt{\newlinechar} (set equal to \texttt{\endlinechar} unless that is equal to 32, before the user \textit{\texttt{\langle setup \rangle}}), then it is split into lines at these characters, then read as if reading multiple lines from a file, ignoring spaces (catcode 10) at the beginning and spaces and tabs (character code 32 or 9) at the end of every line. Otherwise, spaces (and tabs) are retained at both ends of the single-line string, as if it appeared in the middle of a line read from a file.

5 \textbf{Token list conditionals}

\texttt{\tl_if_blank_p:n \tl_if_blank_p:(e|V|o) \tl_if_blank:nTF \tl_if_blank:(e|V|o)TF}

Tests if the \textit{\texttt{\langle token list \rangle}} consists only of blank spaces (\textit{i.e.} contains no item). The test is \texttt{true} if \textit{\texttt{\langle token list \rangle}} is zero or more explicit space characters (explicit tokens with character code 32 and category code 10), and \texttt{false} otherwise.
Tests if the ⟨token list variable⟩ is entirely empty (i.e. contains no tokens at all).

Tests if the ⟨token list⟩ is entirely empty (i.e. contains no tokens at all).

Compares the content of two ⟨token list variables⟩ and is logically true if the two contain the same list of tokens (i.e. identical in both the list of characters they contain and the category codes of those characters). Thus for example

\begin{verbatim}
\tl_set:Nn \l_tmpa_tl { abc }
\tl_set:Nx \l_tmpb_tl { \tl_to_str:n { abc } }
\tl_if_eq:NNTF \l_tmpa_tl \l_tmpb_tl \{ true \} \{ false \}
\end{verbatim}

yields false.

Tests if ⟨token list1⟩ and ⟨token list2⟩ contain the same list of tokens, both in respect of character codes and category codes.

Tests if the ⟨token list⟩ is found in the content of the ⟨tl var⟩. The ⟨token list⟩ cannot contain the tokens {, } or # (more precisely, explicit character tokens with category code 1 (begin-group) or 2 (end-group), and tokens with category code 6).

Tests if ⟨token list2⟩ is found inside ⟨token list1⟩. The ⟨token list2⟩ cannot contain the tokens {, } or # (more precisely, explicit character tokens with category code 1 (begin-group) or 2 (end-group), and tokens with category code 6).

Tests if the ⟨token list⟩ is exactly equal to the special \c_novalue_tl marker. This function is intended to allow construction of flexible document interface structures in which missing optional arguments are detected.
Tests if the content of the \(\langle\text{tl var}\rangle\) consists of a single item, i.e. a single normal token (neither an explicit space character nor a begin-group character) or a single brace group, surrounded by optional spaces on both sides. In other words, such a token list has token count 1 according to \(\text{tl_count:}\text{N}\).

Tests if the \(\langle\text{token list}\rangle\) has exactly one item, i.e. is a single normal token (neither an explicit space character nor a begin-group character) or a single brace group, surrounded by optional spaces on both sides. In other words, such a token list has token count 1 according to \(\text{tl_count:}\text{n}\).

Tests if the token list consists of exactly one token, i.e. is either a single space character or a single “normal” token. Token groups (\{...\}) are not single tokens.

This function compares the \(\langle\text{test token list variable}\rangle\) in turn with each of the \(\langle\text{token list variable cases}\rangle\). If the two are equal (as described for \(\text{tl_if_eq:}\text{NNTF}\)) then the associated \(\langle\text{code}\rangle\) is left in the input stream and other cases are discarded. If any of the cases are matched, the \(\langle\text{true code}\rangle\) is also inserted into the input stream (after the code for the appropriate case), while if none match then the \(\langle\text{false code}\rangle\) is inserted. The function \(\text{tl_case:}\text{Nn}\), which does nothing if there is no match, is also available.

6 Mapping to token lists

All mappings are done at the current group level, i.e. any local assignments made by the \(\langle\text{function}\rangle\) or \(\langle\text{code}\rangle\) discussed below remain in effect after the loop.
\tl_map_function:nN \{⟨token list⟩\} \{⟨function⟩\}

Applies \{⟨function⟩\} to every \{⟨item⟩\} in the \{⟨token list⟩\}. The \{⟨function⟩\} receives one argument for each iteration. This may be a number of tokens if the \{⟨item⟩\} was stored within braces. Hence the \{⟨function⟩\} should anticipate receiving n-type arguments. See also \tl_map_function:nN.

\tl_map_inline:Nn
\tl_map_inline:cn
Updated: 2012-06-29

\tl_map_inline:Nn \tl_map_inline:cn
\tl_map_inline:nn
Updated: 2012-06-29

\tl_map_tokens:Nn \tl_map_tokens:cn \tl_map_tokens:nn
Updated: 2012-06-29

\tl_map_tokens:Nn \tl_map_tokens:cn \tl_map_tokens:nn
New: 2019-09-02

\tl_map_tokens:Nn \tl_my_tl \{\prg_replicate:nn \{2\}\}
expands to twice each item in the \{sequence\}: for each item in \tl_my_tl the function \prg_replicate:nn receives 2 and \{item\} as its two arguments. The function \tl_map_tokens:Nn is typically faster but is not expandable.

\tl_map_variable:NNn
\tl_map_variable:cn
\tl_map_variable:nn
Updated: 2012-06-29

\tl_map_variable:NNn \tl_map_variable:cn \tl_map_variable:nn
Updated: 2012-06-29

\tl_map_variable:NNn \tl_map_variable:cn \tl_map_variable:nn
Updated: 2012-06-29
\tl_map_break: \texttt{\tl_map_break:}

Used to terminate a \texttt{\tl_map...} function before all entries in the \emph{(token list variable)} have been processed. This normally takes place within a conditional statement, for example

\begin{verbatim}
\tl_map_inline:Nn \l_my_tl
{ \str_if_eq:nnT { #1 } { bingo } { \tl_map_break: } % Do something useful
}
\end{verbatim}

See also \texttt{\tl_map_break:n}. Use outside of a \texttt{\tl_map...} scenario leads to low level \TeX{} errors.

\TeX{}hackers note: When the mapping is broken, additional tokens may be inserted before the \emph{(tokens)} are inserted into the input stream. This depends on the design of the mapping function.

\tl_map_break:n \texttt{\tl_map_break:n \{\langle code\rangle\}}

Used to terminate a \texttt{\tl_map...} function before all entries in the \emph{(token list variable)} have been processed, inserting the \emph{(code)} after the mapping has ended. This normally takes place within a conditional statement, for example

\begin{verbatim}
\tl_map_inline:Nn \l_my_tl
{ \str_if_eq:nnT { #1 } { bingo } { \tl_map_break:n \{ <code> \} } % Do something useful
}
\end{verbatim}

Use outside of a \texttt{\tl_map...} scenario leads to low level \TeX{} errors.

\TeX{}hackers note: When the mapping is broken, additional tokens may be inserted before the \emph{(code)} is inserted into the input stream. This depends on the design of the mapping function.
7 Using token lists

\tl_to_str:n \tl_to_str:n \{\textit{token list}\}

Converts the \textit{token list} to a \textit{string}, leaving the resulting character tokens in the input stream. A \textit{string} is a series of tokens with category code 12 \textit{(other)} with the exception of spaces, which retain category code 10 \textit{(space)}. This function requires only a single expansion. Its argument \textit{must} be braced.

\textbf{\TeXhackers note:} This is the $\epsilon$-\TeX primitive $\detokenize$. Converting a \textit{token list} to a \textit{string} yields a concatenation of the string representations of every token in the \textit{token list}. The string representation of a control sequence is

1. an escape character, whose character code is given by the internal parameter $\backslash\texttt{escapechar}$, absent if the $\backslash\texttt{escapechar}$ is negative or greater than the largest character code;
2. the control sequence name, as defined by $\texttt{cs\_to\_str:N}$;
3. a space, unless the control sequence name is a single character whose category at the time of expansion of $\tl_to_str:n$ is not "letter".

The string representation of an explicit character token is that character, doubled in the case of (explicit) macro parameter characters (normally \#). In particular, the string representation of a token list may depend on the category codes in effect when it is evaluated, and the value of the $\backslash\texttt{escapechar}$: for instance $\tl_to_str:n \{\texttt{a}\}$ normally produces the three character “backslash”, “lower-case a”, “space”, but it may also produce a single “lower-case a” if the escape character is negative and a is currently not a letter.

\tl_to_str:N \tl_to_str:N \tl_var

Converts the content of the \textit{tl var} into a series of characters with category code 12 \textit{(other)} with the exception of spaces, which retain category code 10 \textit{(space)}. This \textit{string} is then left in the input stream. For low-level details, see the notes given for $\tl_to_str:n$.

\tl_use:N \tl_use:N \tl_var

Recovers the content of a \textit{tl var} and places it directly in the input stream. An error is raised if the variable does not exist or if it is invalid. Note that it is possible to use a \textit{tl var} directly without an accessor function.

8 Working with the content of token lists

\tl_count:n \tl_count:n \{\textit{tokens}\}

Counts the number of \textit{items} in \textit{tokens} and leaves this information in the input stream. Unbraced tokens count as one element as do each token group \{\ldots\}. This process ignores any unprotected spaces within \textit{tokens}. See also $\tl_count:N$. This function requires three expansions, giving an \textit{integer denotation}. 
\tl_count:N \{tl var\}
Counts the number of token groups in the ⟨tl var⟩ and leaves this information in the input stream. Unbraced tokens count as one element as do each token group (⟨...⟩). This process ignores any unprotected spaces within the ⟨tl var⟩. See also \tl_count:n. This function requires three expansions, giving an ⟨integer denotation⟩.

\tl_count_tokens:n \{tokens\}
Counts the number of TEX tokens in the ⟨tokens⟩ and leaves this information in the input stream. Every token, including spaces and braces, contributes one to the total; thus for instance, the token count of a~{bc} is 6.

\tl_reverse:n \{token list\}
Reverses the order of the ⟨items⟩ in the ⟨token list⟩, so that ⟨item₁⟩ ⟨item₂⟩ ⟨item₃⟩ ... ⟨itemₙ⟩ becomes ⟨itemₙ⟩ ... ⟨item₃⟩ ⟨item₂⟩ ⟨item₁⟩. This process preserves unprotected space within the ⟨token list⟩. Tokens are not reversed within braced token groups, which keep their outer set of braces. In situations where performance is important, consider \tl_reverse_items:n. See also \tl_reverse:N.

\tl_reverse_items:n \{token list\}
Reverses the order of the ⟨items⟩ stored in ⟨tl var⟩, so that ⟨item₁⟩ ⟨item₂⟩ ⟨item₃⟩ ... ⟨itemₙ⟩ becomes ⟨itemₙ⟩ ... ⟨item₃⟩ ⟨item₂⟩ ⟨item₁⟩. This process removes any unprotected space within the ⟨token list⟩. Braced token groups are copied without reversing the order of tokens, and keep the outer set of braces. Items which are initially not braced are copied with braces in the result. In cases where preserving spaces is important, consider the slower function \tl_reverse:n.

\tl_trim_spaces:n \{token list\}
Removes any leading and trailing explicit space characters (explicit tokens with character code 32 and category code 10) from the ⟨token list⟩ and leaves the result in the input stream.

\TeXhackers note: The result is returned within \unexpanded, which means that the token list does not expand further when appearing in an x-type argument expansion.
\tl_trim_spaces_apply:nN \star \tl trim_spaces_apply:oN \star

Removes any leading and trailing explicit space characters (explicit tokens with character code 32 and category code 10) from the \token list and passes the result to the \function as an n-type argument.

\tl_trim_spaces:N \tl_trim_spaces:c \tl_gtrim_spaces:N \tl_gtrim_spaces:c

New: 2011-07-09

\tl_sort:Nn \tl_sort:cn \tl_gsort:Nn \tl_gsort:cn

New: 2017-02-06

\tl_sort:nN \tl_sort:n \tl_gsort:n \tl_gsort:n

New: 2017-02-06

\tl_sort:nN \token list \ conditional

Sorts the items in the \token list, using the \conditional to compare items, and leaves the result in the input stream. The \conditional should have signature :nnTF, and return \true if the two items being compared should be left in the same order, and \false if the items should be swapped. The details of sorting comparison are described in Section 1.

\TeXhackers note: The result is returned within \exp_not:n, which means that the token list does not expand further when appearing in an x-type or e-type argument expansion.

9 The first token from a token list

Functions which deal with either only the very first item (balanced text or single normal token) in a token list, or the remaining tokens.
Leaves in the input stream the first \langle \text{item} \rangle in the \langle \text{token list} \rangle, discarding the rest of the \langle \text{token list} \rangle. All leading explicit space characters (explicit tokens with character code 32 and category code 10) are discarded; for example

\tl_head:n \{ abc \}

and

\tl_head:n \{ - abc \}

both leave a in the input stream. If the “head” is a brace group, rather than a single token, the braces are removed, and so

\tl_head:n \{ - \{ - ab \} c \}

yields ab. A blank \langle \text{token list} \rangle (see \tl_if_blank:nT) results in \tl_head:n leaving nothing in the input stream.

\textbf{\TeX{}hackers note}: The result is returned within \exp_not:n, which means that the token list does not expand further when appearing in an x-type argument expansion.

\tl_tail:n \langle \text{token list} \rangle \{ \} \q_stop

Leaves in the input stream the first \langle \text{item} \rangle in the \langle \text{token list} \rangle, discarding the rest of the \langle \text{token list} \rangle. All leading explicit space characters (explicit tokens with character code 32 and category code 10) are discarded. A blank \langle \text{token list} \rangle (which consists only of space characters) results in a low-level \TeX{} error, which may be avoided by the inclusion of an empty group in the input (as shown), without the need for an explicit test. Alternatively, \tl_if_blank:nF may be used to avoid using the function with a “blank” argument. This function requires only a single expansion, and thus is suitable for use within an o-type expansion. In general, \tl_tail:n should be preferred if the number of expansions is not critical.

\tl_tail:n \langle \text{token list} \rangle \{ \}

Discards all leading explicit space characters (explicit tokens with character code 32 and category code 10) and the first \langle \text{item} \rangle in the \langle \text{token list} \rangle, and leaves the remaining tokens in the input stream. Thus for example

\tl_tail:n \{ a - \{bc\} d \}

and

\tl_tail:n \{ - a - \{bc\} d \}

both leave \{bc\}d in the input stream. A blank \langle \text{token list} \rangle (see \tl_if_blank:nT) results in \tl_tail:n leaving nothing in the input stream.

\textbf{\TeX{}hackers note}: The result is returned within \exp_not:n, which means that the token list does not expand further when appearing in an x-type argument expansion.
Tests if the first \( \langle \text{token} \rangle \) in the \( \langle \text{token list} \rangle \) has the same category code as the \( \langle \text{test token} \rangle \).

In the case where the \( \langle \text{token list} \rangle \) is empty, the test is always \text{false}.

Tests if the first \( \langle \text{token} \rangle \) in the \( \langle \text{token list} \rangle \) has the same character code as the \( \langle \text{test token} \rangle \).

In the case where the \( \langle \text{token list} \rangle \) is empty, the test is always \text{false}.

Tests if the first \( \langle \text{token} \rangle \) in the \( \langle \text{token list} \rangle \) has the same meaning as the \( \langle \text{test token} \rangle \). In the case where \( \langle \text{token list} \rangle \) is empty, the test is always \text{false}.

Tests if the first \( \langle \text{token} \rangle \) in the \( \langle \text{token list} \rangle \) is an explicit begin-group character (with category code 1 and any character code), in other words, if the \( \langle \text{token list} \rangle \) starts with a brace group. In particular, the test is \text{false} if the \( \langle \text{token list} \rangle \) starts with an implicit token such as \texttt{\c_group_begin_token}, or if it is empty. This function is useful to implement actions on token lists on a token by token basis.

Tests if the first \( \langle \text{token} \rangle \) in the \( \langle \text{token list} \rangle \) is a normal \text{N}-type argument. In other words, it is neither an explicit space character (explicit token with character code 32 and category code 10) nor an explicit begin-group character (with category code 1 and any character code). An empty argument yields \text{false}, as it does not have a “normal” first token. This function is useful to implement actions on token lists on a token by token basis.

Tests if the first \( \langle \text{token} \rangle \) in the \( \langle \text{token list} \rangle \) is an explicit space character (explicit token with character code 12 and category code 10). In particular, the test is \text{false} if the \( \langle \text{token list} \rangle \) starts with an implicit token such as \texttt{\c_space_token}, or if it is empty. This function is useful to implement actions on token lists on a token by token basis.
10 Using a single item

\tl_item:nn \{\langle \text{token list} \rangle \} \{\langle \text{integer expression} \rangle \}

Indexing items in the \langle token list \rangle from 1 on the left, this function evaluates the \langle integer expression \rangle and leaves the appropriate item from the \langle token list \rangle in the input stream. If the \langle integer expression \rangle is negative, indexing occurs from the right of the token list, starting at \(-1\) for the right-most item. If the index is out of bounds, then the function expands to nothing.

\textbf{\TeXhackers note:} The result is returned within the \unexpanded primitive (\exp_not:n), which means that the \langle item \rangle does not expand further when appearing in an x-type argument expansion.

\tl_rand_item:N \{\langle \text{tl var} \rangle \}
\tl_rand_item:n \{\langle \text{token list} \rangle \}

Selects a pseudo-random item of the \langle token list \rangle. If the \langle token list \rangle is blank, the result is empty. This is not available in older versions of \TeX.

\textbf{\TeXhackers note:} The result is returned within the \unexpanded primitive (\exp_not:n), which means that the \langle item \rangle does not expand further when appearing in an x-type argument expansion.
Leaves in the input stream the items from the \langle start index \rangle to the \langle end index \rangle inclusive. Spaces and braces are preserved between the items returned (but never at either end of the list). Here \langle start index \rangle and \langle end index \rangle should be \langle integer expressions \rangle. For describing in detail the functions' behavior, let \( m \) and \( n \) be the start and end index respectively. If either is 0, the result is empty. A positive index means ‘start counting from the left end’, and a negative index means ‘from the right end’. Let \( l \) be the count of the token list.

The actual start point is determined as \( M = m \) if \( m > 0 \) and as \( M = l + m + 1 \) if \( m < 0 \). Similarly the actual end point is \( N = n \) if \( n > 0 \) and \( N = l + n + 1 \) if \( n < 0 \). If \( M > N \), the result is empty. Otherwise it consists of all items from position \( M \) to position \( N \) inclusive; for the purpose of this rule, we can imagine that the token list extends at infinity on either side, with void items at positions \( s \) for \( s \leq 0 \) or \( s > l \).

Spaces in between items in the actual range are preserved. Spaces at either end of the token list will be removed anyway (think to the token list being passed to \texttt{\tl_trim_spaces:n} to begin with.

Thus, with \( l = 7 \) as in the examples below, all of the following are equivalent and result in the whole token list

\[
\tl_range:nnn \{ abcd-{e{}}fg \} \{ 1 \} \{ 7 \}
\]
\[
\tl_range:nnn \{ abcd-{e{}}fg \} \{ 1 \} \{ 12 \}
\]
\[
\tl_range:nnn \{ abcd-{e{}}fg \} \{ -7 \} \{ 7 \}
\]
\[
\tl_range:nnn \{ abcd-{e{}}fg \} \{ -12 \} \{ 7 \}
\]

Here are some more interesting examples. The calls

\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ 2 \} \{ 5 \} \}
\]
\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ 2 \} \{ -3 \} \}
\]
\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ -6 \} \{ 5 \} \}
\]
\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ -6 \} \{ -3 \} \}
\]

are all equivalent and will print \texttt{bcd{}{}} on the terminal; similarly

\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ 2 \} \{ 5 \} \}
\]
\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ 2 \} \{ -3 \} \}
\]
\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ -6 \} \{ 5 \} \}
\]
\[
\texttt{\iow_term:x} \{ \tl_range:nnn \{ abcd-{e{}}fg \} \{ -6 \} \{ -3 \} \}
\]

are all equivalent and will print \texttt{bcd \{}{} on the terminal (note the space in the middle).

To the contrary,

\[
\tl_range:nnn \{ abcd-{e{}}f \} \{ 2 \} \{ 4 \}
\]

will discard the space after ‘d’.

If we want to get the items from, say, the third to the last in a token list \texttt{<tl>}, the call is \texttt{\tl_range:nnn \{ <tl> \} \{ 3 \} \{ -1 \}}. Similarly, for discarding the last item, we can do \texttt{\tl_range:nnn \{ <tl> \} \{ 1 \} \{ -2 \}}.

For better performance, see \texttt{\_range_braced:nnn} and \texttt{\_range_unbraced:nnn}.

\textbf{TeXhackers note:} The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp_not:n}), which means that the \texttt{\langle item \rangle} does not expand further when appearing in an \texttt{x}-type argument expansion.
11 Viewing token lists

\tl_show:N \tl_show:N (\tl var)
Displays the content of the \langle \tl var \rangle on the terminal.

\textbf{\TeXhackers note:} This is similar to the \TeX primitive \texttt{\show}, wrapped to a fixed number of characters per line.

\tl_show:n \tl_show:n \{\langle \text{token list} \rangle\}
Displays the \langle \text{token list} \rangle on the terminal.

\textbf{\TeXhackers note:} This is similar to the \TeX primitive \texttt{\showtokens}, wrapped to a fixed number of characters per line.

\tl_log:N \tl_log:N (\tl var)
\tl_log:n \tl_log:n \{\langle \text{token list} \rangle\}
Writes the content of the \langle \tl var \rangle in the log file. See also \texttt{\tl_show:N} which displays the result in the terminal.

12 Constant token lists

\c_empty_tl
Constant that is always empty.

\c_novalue_tl
A marker for the absence of an argument. This constant \texttt{\tl} can safely be typeset (cf. \texttt{\q(nil)}, with the result being \texttt{-NoValue-}. It is important to note that \texttt{\c_novalue_tl} is constructed such that it will \emph{not} match the simple text input \texttt{-NoValue-}, \ie that

\tl_if_eq:VnTF \c_novalue_tl \{ -NoValue- \}

is logically \texttt{false}. The \texttt{\c_novalue_tl} marker is intended for use in creating document-level interfaces, where it serves as an indicator that an (optional) argument was omitted. In particular, it is distinct from a simple empty \texttt{\tl}.

\c_space_tl
An explicit space character contained in a token list (compare this with \texttt{\c_space_token}). For use where an explicit space is required.
13 Scratch token lists

\l_tmpa_tl Scratch token lists for local assignment. These are never used by the kernel code, and so are safe for use with any \TeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\l_tmpb_tl

\g_tmpa_tl Scratch token lists for global assignment. These are never used by the kernel code, and so are safe for use with any \TeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpb_tl
Part VII

The \texttt{l3str} package: Strings

\TeX\ associates each character with a category code: as such, there is no concept of a “string” as commonly understood in many other programming languages. However, there are places where we wish to manipulate token lists while in some sense “ignoring” category codes: this is done by treating token lists as strings in a \TeX\ sense.

A \TeX\ string (and thus an \texttt{expl3} string) is a series of characters which have category code 12 (“other”) with the exception of space characters which have category code 10 (“space”). Thus at a technical level, a \TeX\ string is a token list with the appropriate category codes. In this documentation, these are simply referred to as strings.

String variables are simply specialised token lists, but by convention should be named with the suffix \texttt{...str}. Such variables should contain characters with category code 12 (other), except spaces, which have category code 10 (blank space). All the functions in this module which accept a token list argument first convert it to a string using \texttt{\tl_to_str:n} for internal processing, and do not treat a token list or the corresponding string representation differently.

As a string is a subset of the more general token list, it is sometimes unclear when one should be used over the other. Use a string variable for data that isn’t primarily intended for typesetting and for which a level of protection from unwanted expansion is suitable. This data type simplifies comparison of variables since there are no concerns about expansion of their contents.

The functions \texttt{\cs_to_str:N}, \texttt{\tl_to_str:n}, \texttt{\tl_to_str:N} and \texttt{\token_to_str:N} (and variants) generate strings from the appropriate input: these are documented in \texttt{l3basics}, \texttt{l3tl} and \texttt{l3token}, respectively.

Most expandable functions in this module come in three flavours:

- \texttt{\str_...:N}, which expect a token list or string variable as their argument;
- \texttt{\str_...:n}, taking any token list (or string) as an argument;
- \texttt{\str_..._ignore_spaces:n}, which ignores any space encountered during the operation: these functions are typically faster than those which take care of escaping spaces appropriately.

1 Building strings

\begin{verbatim}
\str_new:N \str_new:c
New: 2015-09-18
\end{verbatim}

\texttt{\str_new:N} \langle str var \rangle

Creates a new \langle str var \rangle or raises an error if the name is already taken. The declaration is global. The \langle str var \rangle is initially empty.

\begin{verbatim}
\str_const:Nn \str_const:(NV|Nx|cn|cV|cx)
New: 2015-09-18
Updated: 2018-07-28
\end{verbatim}

\texttt{\str_const:Nn} \langle str var \rangle {\langle token list \rangle}

Creates a new constant \langle str var \rangle or raises an error if the name is already taken. The value of the \langle str var \rangle is set globally to the \langle token list \rangle, converted to a string.
C.2 Adding data to string variables

\str_set:N \langle str var \rangle \{ \langle token list \rangle \}

Converts the \langle token list \rangle to a \langle string \rangle, and stores the result in \langle str var \rangle.

\str_set:N \langle str var \rangle \langle token list \rangle

Converts the \langle token list \rangle to a \langle string \rangle, and prepends the result to \langle str var \rangle. The current contents of the \langle str var \rangle are not automatically converted to a string.
3 Modifying string variables

\texttt{\texttt{\texttt{str_remove_once:Nn}} \str_remove_once:Nn \{\texttt{str var}\} \{(\texttt{token list})\}}
\texttt{\texttt{\texttt{\texttt{\texttt{\str_gremove_once:Nn}}}} \str_gremove_once:Nn \{\texttt{str var}\} \{(\texttt{token list})\}}

Converts the \texttt{(token list)} to a \texttt{(string)}, and removes the first (leftmost) occurrence of \texttt{(string)} from the \texttt{(str var)}. As this function operates from right to left, the pattern \texttt{(string)} may remain after the removal, for instance,

\texttt{\texttt{\texttt{\str_set:Nn \l_tmpa_str {abbccd} \str_remove_all:Nn \l_tmpa_str \{bc\}}}}

results in \texttt{\l_tmpa_str} containing \texttt{abcd}.
## 4 String conditionals

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\str_if_exist_p:N</td>
<td>Tests whether the (&lt;str\ var&gt;) is currently defined. This does not check that the (&lt;str\ var&gt;) really is a string.</td>
</tr>
<tr>
<td>\str_if_exist_p:C</td>
<td></td>
</tr>
<tr>
<td>\str_if_exist_p:NTF</td>
<td></td>
</tr>
<tr>
<td>\str_if_exist_p:c</td>
<td></td>
</tr>
</tbody>
</table>

New: 2015-09-18

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\str_if_empty_p:N</td>
<td>Tests if the (&lt;string variable&gt;) is entirely empty (i.e. contains no characters at all).</td>
</tr>
<tr>
<td>\str_if_empty_p:C</td>
<td></td>
</tr>
<tr>
<td>\str_if_empty_p:NTF</td>
<td></td>
</tr>
<tr>
<td>\str_if_empty_p:c</td>
<td></td>
</tr>
</tbody>
</table>

New: 2015-09-18

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\str_if_eq_p:NN</td>
<td>Compares the content of two (&lt;str variables&gt;) and is logically true if the two contain the same characters in the same order.</td>
</tr>
<tr>
<td>\str_if_eq_p:(Nc</td>
<td>cN</td>
</tr>
<tr>
<td>\str_if_eq:p:NN</td>
<td></td>
</tr>
<tr>
<td>\str_if_eq:p:cn</td>
<td></td>
</tr>
</tbody>
</table>

New: 2019-06-18

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\str_if_eq:p:nn</td>
<td>Compares the two (&lt;token lists&gt;) on a character by character basis (namely after converting them to strings), and is true if the two (&lt;strings&gt;) contain the same characters in the same order. Thus for example (&lt;\text{abc}&gt;) (&lt;\text{abc}&gt;) is logically true.</td>
</tr>
<tr>
<td>\str_if_eq:p:Vn</td>
<td></td>
</tr>
<tr>
<td>\str_if_eq:p:on</td>
<td></td>
</tr>
<tr>
<td>\str_if_eq:p:no</td>
<td></td>
</tr>
</tbody>
</table>

Updated: 2018-06-18

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\str_if_in:NnTF</td>
<td>Converts both (&lt;token lists&gt;) to (&lt;strings&gt;) and tests whether (&lt;string2&gt;) is found inside (&lt;string1&gt;).</td>
</tr>
<tr>
<td>\str_if_in:nTF</td>
<td></td>
</tr>
</tbody>
</table>

New: 2017-10-08
Compares the ⟨test string⟩ in turn with each of the ⟨string cases⟩ (all token lists are converted to strings). If the two are equal (as described for \str_if_eq:nnTF) then the associated ⟨code⟩ is left in the input stream and other cases are discarded. If any of the cases are matched, the ⟨true code⟩ is also inserted into the input stream (after the code for the appropriate case), while if none match then the ⟨false code⟩ is inserted. The function \str_case:nn, which does nothing if there is no match, is also available.

Compares the full expansion of the ⟨test string⟩ in turn with the full expansion of the ⟨string cases⟩ (all token lists are converted to strings). If the two full expansions are equal (as described for \str_if_eq:nnTF) then the associated ⟨code⟩ is left in the input stream and other cases are discarded. If any of the cases are matched, the ⟨true code⟩ is also inserted into the input stream (after the code for the appropriate case), while if none match then the ⟨false code⟩ is inserted. The function \str_case:nn, which does nothing if there is no match, is also available.

All mappings are done at the current group level, \textit{i.e.} any local assignments made by the ⟨function⟩ or ⟨code⟩ discussed below remain in effect after the loop.

\section{Mapping to strings}

All mappings are done at the current group level, \textit{i.e.} any local assignments made by the ⟨function⟩ or ⟨code⟩ discussed below remain in effect after the loop.
\str_map_inline:Nn \str_map_inline:cn
\textit{New: 2017-11-14}

Applies the \langle inline function \rangle to every \langle character \rangle in the \langle str var \rangle including spaces. The \langle inline function \rangle should consist of code which receives the \langle character \rangle as \#1. See also \str_map_function:NN.

\str_map_inline:nn
\textit{New: 2017-11-14}

Converts the \langle token list \rangle to a \langle string \rangle then applies the \langle inline function \rangle to every \langle character \rangle in the \langle string \rangle including spaces. The \langle inline function \rangle should consist of code which receives the \langle character \rangle as \#1. See also \str_map_function:NN.

\str_map_variable:NNn \str_map_variable:cNn
\textit{New: 2017-11-14}

Stores each \langle character \rangle of the \langle string \rangle (including spaces) in turn in the (string or token list) \langle variable \rangle and applies the \langle code \rangle. The \langle code \rangle will usually make use of the \langle variable \rangle, but this is not enforced. The assignments to the \langle variable \rangle are local. Its value after the loop is the last \langle character \rangle in the \langle string \rangle, or its original value if the \langle string \rangle is empty. See also \str_map_inline:Nn.

\str_map_variable:nn
\textit{New: 2017-11-14}

Converts the \langle token list \rangle to a \langle string \rangle then stores each \langle character \rangle in the \langle string \rangle (including spaces) in turn in the (string or token list) \langle variable \rangle and applies the \langle code \rangle. The \langle code \rangle will usually make use of the \langle variable \rangle, but this is not enforced. The assignments to the \langle variable \rangle are local. Its value after the loop is the last \langle character \rangle in the \langle string \rangle, or its original value if the \langle string \rangle is empty. See also \str_map_inline:Nn.

\str_map_break:
\textit{New: 2017-10-08}

Used to terminate a \str_map_... function before all characters in the \langle string \rangle have been processed. This normally takes place within a conditional statement, for example

\begin{verbatim}
\str_map_inline:Nn \l_my_str
{ \str_if_eq:nnT { #1 } { bingo } { \str_map_break: }
  % Do something useful
}
\end{verbatim}

See also \str_map_break:n. Use outside of a \str_map_... scenario leads to low level \TeX errors.

\textit{TeXhackers note:} When the mapping is broken, additional tokens may be inserted before continuing with the code that follows the loop. This depends on the design of the mapping function.
\str_map_break:n \langle code \rangle

New: 2017-10-08

\str_map_break:n \langle code \rangle

Used to terminate a \str_map_:... function before all characters in the \langle string \rangle have been processed, inserting the \langle code \rangle after the mapping has ended. This normally takes place within a conditional statement, for example

\str_map_inline:Nn \l_my_str
{\str_if_eq:nnT { #1 } { bingo } { \str_map_break:n { <code> } }}
% Do something useful}

Use outside of a \str_map_:... scenario leads to low level \TeX errors.

\TeXhackers note: When the mapping is broken, additional tokens may be inserted before the \langle code \rangle is inserted into the input stream. This depends on the design of the mapping function.

6  Working with the content of strings

\str_use:N \str_use:c

\str_use:N \langle str var \rangle

Recovers the content of a \langle str var \rangle and places it directly in the input stream. An error is raised if the variable does not exist or if it is invalid. Note that it is possible to use a \langle str \rangle directly without an accessor function.

\str_count:n \str_count:N \str_count:c \str_count:n \str_count_ignore_spaces:n

\str_count_spaces:n \str_count_spaces:N \str_count_spaces:c \str_count_spaces:n

Leaves in the input stream the number of characters in the string representation of \langle token list \rangle, as an integer denotation. The functions differ in their treatment of spaces. In the case of \str_count:N and \str_count:n, all characters including spaces are counted. The \str_count_ignore_spaces:n function leaves the number of non-space characters in the input stream.

\str_count_spaces:n \str_count_spaces:n \str_count_spaces:n

Leaves in the input stream the number of space characters in the string representation of \langle token list \rangle, as an integer denotation. Of course, this function has no \_ignore_spaces variant.
Converts the \textit{token list} into a \textit{string}. The first character in the \textit{string} is then left in the input stream, with category code “other”. The functions differ if the first character is a space: \texttt{\textbackslash str\_head:N} and \texttt{\textbackslash str\_head:n} return a space token with category code 10 (blank space), while the \texttt{\textbackslash str\_head\_ignore\_spaces:n} function ignores this space character and leaves the first non-space character in the input stream. If the \textit{string} is empty (or only contains spaces in the case of the \texttt{\_ignore\_spaces} function), then nothing is left on the input stream.

Converts the \textit{token list} to a \textit{string}, removes the first character, and leaves the remaining characters (if any) in the input stream, with category codes 12 and 10 (for spaces). The functions differ in the case where the first character is a space: \texttt{\textbackslash str\_tail:N} and \texttt{\textbackslash str\_tail:n} only trim that space, while \texttt{\textbackslash str\_tail\_ignore\_spaces:n} removes the first non-space character and any space before it. If the \textit{token list} is empty (or blank in the case of the \texttt{\_ignore\_spaces} variant), then nothing is left on the input stream.

Converts the \textit{token list} to a \textit{string}, and leaves in the input stream the character in position \textit{integer expression} of the \textit{string}, starting at 1 for the first (left-most) character. In the case of \texttt{\textbackslash str\_item:Nn} and \texttt{\textbackslash str\_item:nn}, all characters including spaces are taken into account. The \texttt{\textbackslash str\_item\_ignore\_spaces:nn} function skips spaces when counting characters. If the \textit{integer expression} is negative, characters are counted from the end of the \textit{string}. Hence, $-1$ is the right-most character, \textit{etc.}
Converting the \langle token list \rangle to a \langle string \rangle, and leaves in the input stream the characters from the \langle start index \rangle to the \langle end index \rangle inclusive. Spaces are preserved and counted as items (contrast this with \texttt{\str\_range:nnn} where spaces are not counted as items and are possibly discarded from the output).

Here \langle start index \rangle and \langle end index \rangle should be integer denotations. For describing in detail the functions’ behavior, let \( m \) and \( n \) be the start and end index respectively. If either is \( 0 \), the result is empty. A positive index means ‘start counting from the left end’, a negative index means ‘start counting from the right end’. Let \( l \) be the count of the token list.

The actual start point is determined as \( M = m \) if \( m > 0 \) and as \( M = l + m + 1 \) if \( m < 0 \). Similarly the actual end point is \( N = n \) if \( n > 0 \) and \( N = l + n + 1 \) if \( n < 0 \). If \( M > N \), the result is empty. Otherwise it consists of all items from position \( M \) to position \( N \) inclusive; for the purpose of this rule, we can imagine that the token list extends at infinity on either side, with void items at positions \( s \) for \( s \leq 0 \) or \( s > l \). For instance,

\begin{verbatim}
\iow_term:x { \str\_range:nnn { abcdef } { 2 } { 5 } }
\iow_term:x { \str\_range:nnn { abcdef } { -4 } { -1 } }
\iow_term:x { \str\_range:nnn { abcdef } { -2 } { -1 } }
\iow_term:x { \str\_range:nnn { abcdef } { 0 } { -1 } }
\end{verbatim}

prints \texttt{bcde}, \texttt{cdef}, \texttt{ef}, and an empty line to the terminal. The \langle start index \rangle must always be smaller than or equal to the \langle end index \rangle: if this is not the case then no output is generated. Thus

\begin{verbatim}
\iow_term:x { \str\_range:nnn { abcdef } { 5 } { 2 } }
\iow_term:x { \str\_range:nnn { abcdef } { -1 } { -4 } }
\end{verbatim}

both yield empty strings.

The behavior of \texttt{\str\_range\_ignore\_spaces:nnn} is similar, but spaces are removed before starting the job. The input

\begin{verbatim}
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abcdefg } { 2 } { 5 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abcdefg } { 2 } { -3 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abcdefg } { -6 } { 5 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abcdefg } { -6 } { -3 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abc-efg } { 2 } { 5 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abc-efg } { 2 } { -3 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abc-efg } { -6 } { 5 } }
\iow_term:x { \str\_range\_ignore\_spaces:nnn { abc-efg } { -6 } { -3 } }
\end{verbatim}

63
will print four instances of bcde, four instances of bc e and eight instances of bcde.

7 String manipulation

\str_lowercase:n {{tokens}} \str_uppercase:n {{tokens}}

Converts the input (tokens) to their string representation, as described for \tl_to_-str:n, and then to the lower or upper case representation using a one-to-one mapping as described by the Unicode Consortium file UnicodeData.txt.

These functions are intended for case changing programmatic data in places where upper/lower case distinctions are meaningful. One example would be automatically generating a function name from user input where some case changing is needed. In this situation the input is programmatic, not textual, case does have meaning and a language-independent one-to-one mapping is appropriate. For example

\cs_new_protected:Npn \myfunc:nn #1#2
{\cs_set_protected:cpn
 { user
 \str_uppercase:f { \tl_head:n {#1} } \str_lowercase:f { \tl_tail:n {#1} } }
 { #2 } }

would be used to generate a function with an auto-generated name consisting of the upper case equivalent of the supplied name followed by the lower case equivalent of the rest of the input.

These functions should not be used for

- Caseless comparisons: use \str_foldcase:n for this situation (case folding is distinct from lower casing).
- Case changing text for typesetting: see the \text_lowercase:n(n), \text_uppercase:n(n) and \text_titlecase:n(n) functions which correctly deal with context-dependence and other factors appropriate to text case changing.

\textbf{\TeX}hackers note: As with all expl3 functions, the input supported by \str_foldcase:n is engine-native characters which are or interoperate with UTF-8. As such, when used with pdf\TeX only the Latin alphabet characters A–Z are case-folded (i.e. the ASCII range which coincides with UTF-8). Full UTF-8 support is available with both X\TeX and Lua\TeX.
\str_foldcase:n \{\langle tokens \rangle\}

Converts the input \{\langle tokens \rangle\} to their string representation, as described for \tl_to_str:n, and then folds the case of the resulting \langle string \rangle to remove case information. The result of this process is left in the input stream.

String folding is a process used for material such as identifiers rather than for “text”. The folding provided by \str_foldcase:n follows the mappings provided by the Unicode Consortium, who state:

Case folding is primarily used for caseless comparison of text, such as identifiers in a computer program, rather than actual text transformation. Case folding in Unicode is based on the lowercase mapping, but includes additional changes to the source text to help make it language-insensitive and consistent. As a result, case-folded text should be used solely for internal processing and generally should not be stored or displayed to the end user.

The folding approach implemented by \str_foldcase:n follows the “full” scheme defined by the Unicode Consortium (e.g. SSfolds to SS). As case-folding is a language-insensitive process, there is no special treatment of Turkic input (i.e. I always folds to i and not to І).

\TeXhackers note: As with all expl3 functions, the input supported by \str_foldcase:n is engine-native characters which are or interoperate with UTF-8. As such, when used with pdf\TeX only the Latin alphabet characters A–Z are case-folded (i.e. the ASCII range which coincides with UTF-8). Full UTF-8 support is available with both X\TeX and Lua\TeX, subject only to the fact that X\TeX in particular has issues with characters of code above hexadecimal 0xFFFF when interacting with \tl_to_str:n.

8 Viewing strings

\str_show:N \str_show:c \str_show:n
\New: 2015-09-18

Displays the content of the \langle str var \rangle on the terminal.

\str_log:N \str_log:c \str_log:n
\New: 2019-02-15

Writes the content of the \langle str var \rangle in the log file.
9 Constant token lists

\c_\_ampersand_str
\c_\_at_str
\c\_backslash_str
\c\_left_brace_str
\c\_right_brace_str
\c\_circumflex_str
\c\_colon_str
\c\_dollar_str
\c\_hash_str
\c\_percent_str
\c\_tilde_str
\c\_underscore_str

New: 2015-09-19

Constant strings, containing a single character token, with category code 12.

10 Scratch strings

\l\tmpa_str
\l\tmpb_str

Scratch strings for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g\tmpa_str
\g\tmpb_str

Scratch strings for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
Part VIII

The l3str-convert package: string encoding conversions

1 Encoding and escaping schemes

Traditionally, string encodings only specify how strings of characters should be stored as bytes. However, the resulting lists of bytes are often to be used in contexts where only a restricted subset of bytes are permitted (e.g., PDF string objects, URLs). Hence, storing a string of characters is done in two steps.

- The code points (“character codes”) are expressed as bytes following a given “encoding”. This can be utf-16, ISO 8859-1, etc. See Table 1 for a list of supported encodings.\footnote{Encodings and escapings will be added as they are requested.}

- Bytes are translated to TEX tokens through a given “escaping”. Those are defined for the most part by the pdf file format. See Table 2 for a list of escaping methods supported.\footnote{}

2 Conversion functions

\begin{verbatim}
\str_set_convert:Nnnn \str_set_convert:Nnnn (str var) {⟨string⟩} {⟨name 1⟩} {⟨name 2⟩}
\str_gset_convert:Nnnn
\end{verbatim}

This function converts the ⟨string⟩ from the encoding given by ⟨name 1⟩ to the encoding given by ⟨name 2⟩, and stores the result in the ⟨str var⟩. Each ⟨name⟩ can have the form ⟨encoding⟩ or ⟨encoding⟩/⟨escaping⟩, where the possible values of ⟨encoding⟩ and ⟨escaping⟩ are given in Tables 1 and 2, respectively. The default escaping is to input and output bytes directly. The special case of an empty ⟨name⟩ indicates the use of “native” strings, 8-bit for pdfTEX, and Unicode strings for the other two engines.

For example,

\begin{verbatim}
\str_set_convert:Nnnn \l_foo_str { Hello! } { } { utf16/hex }
\end{verbatim}

results in the variable \l_foo_str holding the string FEFF00480065006C006C0021. This is obtained by converting each character in the (native) string Hello! to the UTF-16 encoding, and expressing each byte as a pair of hexadecimal digits. Note the presence of a (big-endian) byte order mark “FEFF, which can be avoided by specifying the encoding utf16be/hex.

An error is raised if the ⟨string⟩ is not valid according to the ⟨escaping 1⟩ and ⟨encoding 1⟩, or if it cannot be reencoded in the ⟨encoding 2⟩ and ⟨escaping 2⟩ (for instance, if a character does not exist in the ⟨encoding 2⟩). Erroneous input is replaced by the Unicode replacement character “FFFD, and characters which cannot be reencoded are replaced by either the replacement character “FFFD if it exists in the ⟨encoding 2⟩, or an encoding-specific replacement character, or the question mark character.
Table 1: Supported encodings. Non-alphanumeric characters are ignored, and capital letters are lower-cased before searching for the encoding in this list.

<table>
<thead>
<tr>
<th>Encoding</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>utf8</td>
<td>UTF-8</td>
</tr>
<tr>
<td>utf16</td>
<td>UTF-16, with byte-order mark</td>
</tr>
<tr>
<td>utf16be</td>
<td>UTF-16, big-endian</td>
</tr>
<tr>
<td>utf16le</td>
<td>UTF-16, little-endian</td>
</tr>
<tr>
<td>utf32</td>
<td>UTF-32, with byte-order mark</td>
</tr>
<tr>
<td>utf32be</td>
<td>UTF-32, big-endian</td>
</tr>
<tr>
<td>utf32le</td>
<td>UTF-32, little-endian</td>
</tr>
<tr>
<td>iso88591</td>
<td>ISO 8859-1</td>
</tr>
<tr>
<td>iso88592</td>
<td>ISO 8859-2</td>
</tr>
<tr>
<td>iso88593</td>
<td>ISO 8859-3</td>
</tr>
<tr>
<td>iso88594</td>
<td>ISO 8859-4</td>
</tr>
<tr>
<td>iso88595</td>
<td>ISO 8859-5</td>
</tr>
<tr>
<td>iso88596</td>
<td>ISO 8859-6</td>
</tr>
<tr>
<td>iso88597</td>
<td>ISO 8859-7</td>
</tr>
<tr>
<td>iso88598</td>
<td>ISO 8859-8</td>
</tr>
<tr>
<td>iso88599</td>
<td>ISO 8859-9</td>
</tr>
<tr>
<td>iso885910</td>
<td>ISO 8859-10</td>
</tr>
<tr>
<td>iso885911</td>
<td>ISO 8859-11</td>
</tr>
<tr>
<td>iso885913</td>
<td>ISO 8859-13</td>
</tr>
<tr>
<td>iso885914</td>
<td>ISO 8859-14</td>
</tr>
<tr>
<td>iso885915</td>
<td>ISO 8859-15</td>
</tr>
<tr>
<td>iso885916</td>
<td>ISO 8859-16</td>
</tr>
<tr>
<td>clist</td>
<td>comma-list of integers</td>
</tr>
<tr>
<td>(empty)</td>
<td>native (Unicode) string</td>
</tr>
</tbody>
</table>

Table 2: Supported escapings. Non-alphanumeric characters are ignored, and capital letters are lower-cased before searching for the escaping in this list.

<table>
<thead>
<tr>
<th>Escaping</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bytes, or empty</td>
<td>arbitrary bytes</td>
</tr>
<tr>
<td>hex, hexadecimal</td>
<td>byte = two hexadecimal digits</td>
</tr>
<tr>
<td>name</td>
<td>see \pdfescapename</td>
</tr>
<tr>
<td>string</td>
<td>see \pdfescapestring</td>
</tr>
<tr>
<td>url</td>
<td>encoding used in URLs</td>
</tr>
</tbody>
</table>
\str_set_convert:NnnnTF \str_gset_convert:NnnnTF \str_set_convert:Nnnn \langle string \rangle \{ \langle name 1 \rangle \} \{ \langle name 2 \rangle \} \{ \langle true code \rangle \} \{ \langle false code \rangle \}

As \str_set_convert:Nnnn, converts the \langle string \rangle from the encoding given by \langle name 1 \rangle to the encoding given by \langle name 2 \rangle, and assigns the result to \langle str var \rangle. Contrarily to \str_set_convert:Nnnn, the conditional variant does not raise errors in case the \langle string \rangle is not valid according to the \langle name 1 \rangle encoding, or cannot be expressed in the \langle name 2 \rangle encoding. Instead, the \langle false code \rangle is performed.

3 Creating 8-bit mappings

\str_declare_eight_bit_encoding:nnn \str_declare_eight_bit_encoding:nnn \langle name \rangle \{ \langle mapping \rangle \} \{ \langle missing \rangle \}

Declares the encoding \langle name \rangle to map bytes to Unicode characters according to the \langle mapping \rangle, and map those bytes which are not mentioned in the \langle mapping \rangle either to the replacement character (if they appear in \langle missing \rangle), or to themselves.

4 Possibilities, and things to do

Encoding/escaping-related tasks.

- In Xe\TeX/Lua\TeX, would it be better to use the ^^^^... approach to build a string from a given list of character codes? Namely, within a group, assign 0-9a-f and all characters we want to category “other”, then assign ^ the category superscript, and use \scantokens.
- Change \str_set_convert:Nnnn to expand its last two arguments.
- Describe the internal format in the code comments. Refuse code points in [‘D800, ‘DFFF] in the internal representation?
- Add documentation about each encoding and escaping method, and add examples.
- The \texttt{hex} unescaping should raise an error for odd-token count strings.
- Decide what bytes should be escaped in the \url escaping. Perhaps the characters !’(*)+-./0123456789_ are safe, and all other characters should be escaped?
- Automate generation of 8-bit mapping files.
- Change the framework for 8-bit encodings: for decoding from 8-bit to Unicode, use 256 integer registers; for encoding, use a tree-box.
- More encodings (see Heiko’s stringenc). CESU?
- More escapings: ascii85, shell escapes, lua escapes, etc.?
Part IX
The \texttt{l3quark} package

Quarks

Two special types of constants in \LaTeX{} are “quarks” and “scan marks”. By convention all constants of type quark start out with \texttt{\q_}, and scan marks start with \texttt{\s_}.

1 Quarks

Quarks are control sequences that expand to themselves and should therefore \textit{never} be executed directly in the code. This would result in an endless loop!

They are meant to be used as delimiter in weird functions, the most common use case being the `stop token’ (\textit{i.e.} \texttt{\q_stop}). For example, when writing a macro to parse a user-defined date

\begin{verbatim}
\date_parse:n {19/June/1981}
\end{verbatim}

one might write a command such as

\begin{verbatim}
\cs_new:Npn \date_parse:n #1 { \date_parse_aux:w #1 \q_stop }
\cs_new:Npn \date_parse_aux:w #1 / #2 / #3 \q_stop
{ <do something with the date> }
\end{verbatim}

Quarks are sometimes also used as error return values for functions that receive erroneous input. For example, in the function \texttt{\prop_get:NnN} to retrieve a value stored in some key of a property list, if the key does not exist then the return value is the quark \texttt{\q_no_value}. As mentioned above, such quarks are extremely fragile and it is imperative when using such functions that code is carefully written to check for pathological cases to avoid leakage of a quark into an uncontrolled environment.

Quarks also permit the following ingenious trick when parsing tokens: when you pick up a token in a temporary variable and you want to know whether you have picked up a particular quark, all you have to do is compare the temporary variable to the quark using \texttt{\tl_if_eq:NNTF}. A set of special quark testing functions is set up below. All the quark testing functions are expandable although the ones testing only single tokens are much faster. An example of the quark testing functions and their use in recursion can be seen in the implementation of \texttt{\clist_map_function:NN}.

2 Defining quarks

\texttt{\quark_new:N} \langle quark \rangle

Creates a new \langle quark \rangle which expands only to \langle quark \rangle. The \langle quark \rangle is defined globally, and an error message is raised if the name was already taken.

\texttt{\q_stop}

Used as a marker for delimited arguments, such as

\begin{verbatim}
\cs_set:Npn \tmp:w #1#2 \q_stop \{#1}
\end{verbatim}
- \texttt{q.mark} Used as a marker for delimited arguments when \texttt{q.stop} is already in use.

- \texttt{q.nil} Quark to mark a null value in structured variables or functions. Used as an end delimiter when this may itself need to be tested (in contrast to \texttt{q.stop}, which is only ever used as a delimiter).

- \texttt{q.no_value} A canonical value for a missing value, when one is requested from a data structure. This is therefore used as a “return” value by functions such as \texttt{prop.get:NnN} if there is no data to return.

### 3 Quark tests

The method used to define quarks means that the single token (N) tests are faster than the multi-token (n) tests. The latter should therefore only be used when the argument can definitely take more than a single token.

\begin{align*}
\texttt{quark\_if\_nil\_p:N} \ast & \quad \texttt{quark\_if\_nil\_p:N} \langle \text{token} \rangle \\
\texttt{quark\_if\_nil\_nTF} \ast & \quad \texttt{quark\_if\_nil\_nTF} \langle \text{token} \rangle \{\langle \text{true code} \rangle\} \{\langle \text{false code} \rangle\}
\end{align*}

Tests if the \langle token \rangle is equal to \texttt{q.nil}.

\begin{align*}
\texttt{quark\_if\_nil\_p:n} \ast & \quad \texttt{quark\_if\_nil\_p:n} \langle \text{token list} \rangle \\
\texttt{quark\_if\_nil\_nTF} \ast & \quad \texttt{quark\_if\_nil\_nTF} \langle \text{token list} \rangle \{\langle \text{true code} \rangle\} \{\langle \text{false code} \rangle\}
\end{align*}

Tests if the \langle token list \rangle contains only \texttt{q.nil} (distinct from \langle token list \rangle being empty or containing \texttt{q.nil} plus one or more other tokens).

\begin{align*}
\texttt{quark\_if\_no\_value\_p:N} \ast & \quad \texttt{quark\_if\_no\_value\_p:N} \langle \text{token} \rangle \\
\texttt{quark\_if\_no\_value\_p:c} \ast & \quad \texttt{quark\_if\_no\_value\_p:c} \langle \text{token} \rangle \\
\texttt{quark\_if\_no\_value\_nTF} \ast & \quad \texttt{quark\_if\_no\_value\_nTF} \langle \text{token} \rangle \{\langle \text{true code} \rangle\} \{\langle \text{false code} \rangle\}
\end{align*}

Tests if the \langle token \rangle is equal to \texttt{q.no_value}.

\begin{align*}
\texttt{quark\_if\_no\_value\_p:n} \ast & \quad \texttt{quark\_if\_no\_value\_p:n} \langle \text{token list} \rangle \\
\texttt{quark\_if\_no\_value\_nTF} \ast & \quad \texttt{quark\_if\_no\_value\_nTF} \langle \text{token list} \rangle \{\langle \text{true code} \rangle\} \{\langle \text{false code} \rangle\}
\end{align*}

Tests if the \langle token list \rangle contains only \texttt{q.no_value} (distinct from \langle token list \rangle being empty or containing \texttt{q.no_value} plus one or more other tokens).

### 4 Recursion

This module provides a uniform interface to intercepting and terminating loops as when one is doing tail recursion. The building blocks follow below and an example is shown in Section 5.

\texttt{q.recursion\_tail} This quark is appended to the data structure in question and appears as a real element there. This means it gets any list separators around it.
This quark is added after the data structure. Its purpose is to make it possible to terminate the recursion at any point easily.

\quark_if_recursion_tail_stop:N \quark_if_recursion_tail_stop:N \langle \text{token} \rangle \quark_if_recursion_tail_stop:N \star 

Tests if \langle \text{token} \rangle contains only the marker \q_recursion_tail, and if so uses \use\_none\_delimit\_by\_q\_recursion\_stop:w to terminate the recursion that this belongs to. The recursion input must include the marker tokens \q_recursion_tail and \q\_recursion\_stop as the last two items.

\quark_if_recursion_tail_stop:n \langle \text{token list} \rangle \quark_if_recursion_tail_stop:n \star 

Tests if the \langle \text{token list} \rangle contains only \q_recursion_tail, and if so uses \use\_none\_delimit\_by\_q\_recursion\_stop:w to terminate the recursion that this belongs to. The recursion input must include the marker tokens \q_recursion_tail and \q\_recursion\_stop as the last two items.

\quark_if_recursion_tail_stop_do:Nn \langle \text{token} \rangle \{ \langle \text{insertion} \rangle \} \quark_if_recursion_tail_stop_do:Nn \star 

Tests if \langle \text{token} \rangle contains only the marker \q_recursion_tail, and if so uses \use\_i\_delimit\_by\_q\_recursion\_stop:w to terminate the recursion that this belongs to. The recursion input must include the marker tokens \q_recursion_tail and \q\_recursion\_stop as the last two items. The \langle \text{insertion} \rangle code is then added to the input stream after the recursion has ended.

\quark_if_recursion_tail_break:nN \langle \text{token list} \rangle \langle \text{type} \rangle\_map\_break: \quark_if_recursion_tail_break:NN \star 

Tests if \langle \text{token list} \rangle contains only \q_recursion_tail, and if so terminates the recursion using \langle \text{type} \rangle\_map\_break:. The recursion end should be marked by \prg_break_point:Nn \langle \text{type} \rangle\_map\_break:.

5 An example of recursion with quarks

Quarks are mainly used internally in the expl3 code to define recursion functions such as \tl_map_inline:nn and so on. Here is a small example to demonstrate how to
use quarks in this fashion. We shall define a command called \my_map_dbl:nn which takes a token list and applies an operation to every pair of tokens. For example, \my_map_dbl:nn \{abcd\} \{[-#1--#2--]\} would produce “[-a-b-] [-c-d-] “. Using quarks to define such functions simplifies their logic and ensures robustness in many cases.

Here’s the definition of \my_map_dbl:nn. First of all, define the function that does the processing based on the inline function argument #2. Then initiate the recursion using an internal function. The token list #1 is terminated using \q_recursion_tail, with delimiters according to the type of recursion (here a pair of \q_recursion_tail), concluding with \q_recursion_stop. These quarks are used to mark the end of the token list being operated upon.

\cs_new:Npn \my_map_dbl:nn #1#2
{\cs_set:Npn \__my_map_dbl_fn:nn ##1 ##2 {#2}
 \__my_map_dbl:nn #1 \q_recursion_tail \q_recursion_tail \q_recursion_stop}

The definition of the internal recursion function follows. First check if either of the input tokens are the termination quarks. Then, if not, apply the inline function to the two arguments.

\cs_new:Nn \__my_map_dbl:nn
{\quark_if_recursion_tail_stop:n {#1}
 \quark_if_recursion_tail_stop:n {#2}
 \__my_map_dbl_fn:nn {#1} {#2} }

Finally, recurse:

\__my_map_dbl:nn
}

Note that contrarily to \LaTeX built-in mapping functions, this mapping function cannot be nested, since the second map would overwrite the definition of \__my_map_dbl_fn:nn.

6 Scan marks

Scan marks are control sequences set equal to \scan_stop:, hence never expand in an expansion context and are (largely) invisible if they are encountered in a typesetting context.

Like quarks, they can be used as delimiters in weird functions and are often safer to use for this purpose. Since they are harmless when executed by \TeX in non-expandable contexts, they can be used to mark the end of a set of instructions. This allows to skip to that point if the end of the instructions should not be performed (see \l3regex).

\scan_new:N \scan_stop: \langle \text{scan mark} \rangle

Creates a new \langle \text{scan mark} \rangle which is set equal to \scan_stop:. The \langle \text{scan mark} \rangle is defined globally, and an error message is raised if the name was already taken by another scan mark.

73
\s_stop

Used at the end of a set of instructions, as a marker that can be jumped to using \use_none_delimit_by_s_stop:w.

\use_none_delimit_by_s_stop:w *
\use_none_delimit_by_s_stop:w (tokens) \s_stop

Removes the \textit{(tokens)} and \s_stop from the input stream. This leads to a low-level TeX error if \s_stop is absent.
Part X
The \texttt{l3seq} package
Sequences and stacks

\LaTeXe{} implements a “sequence” data type, which contain an ordered list of entries which may contain any \textit{(balanced text)}. It is possible to map functions to sequences such that the function is applied to every item in the sequence.

Sequences are also used to implement stack functions in \LaTeXe{}. This is achieved using a number of dedicated stack functions.

1 Creating and initialising sequences

\begin{verbatim}
\seq_new:N \seq_new:c
\seq_clear:N \seq_clear:c \seq_gclear:N \seq_gclear:c
\seq_clear_new:N \seq_clear_new:c \seq_gclear_new:N \seq_gclear_new:c
\seq_set_eq:NN \seq_set_eq:\ (cN|Nc|cc)
\seq_gset_eq:NN \seq_gset_eq:\ (cN|Nc|cc)
\seq_set_from_clist:NN \seq_set_from_clist:(cN|Nc|cc)
\seq_set_from_clist:Nn \seq_set_from_clist:cn
\seq_gset_from_clist:NN \seq_gset_from_clist:(cN|Nc|cc)
\seq_gset_from_clist:Nn \seq_gset_from_clist:cn
\end{verbatim}

\texttt{\seq_new:N \langle sequence\rangle} \hspace{2cm} \texttt{\seq_new:c \langle sequence\rangle}

Creates a new \textit{(sequence)} or raises an error if the name is already taken. The declaration is global. The \textit{(sequence)} initially contains no items.

\texttt{\seq_clear:N \langle sequence\rangle} \hspace{2cm} \texttt{\seq_clear:c \langle sequence\rangle} \hspace{2cm} \texttt{\seq_gclear:N \langle sequence\rangle} \hspace{2cm} \texttt{\seq_gclear:c \langle sequence\rangle}

Clears all items from the \textit{(sequence)}.

\texttt{\seq_clear_new:N \langle sequence\rangle} \hspace{2cm} \texttt{\seq_clear_new:c \langle sequence\rangle} \hspace{2cm} \texttt{\seq_gclear_new:N \langle sequence\rangle} \hspace{2cm} \texttt{\seq_gclear_new:c \langle sequence\rangle}

Ensures that the \textit{(sequence)} exists globally by applying \texttt{\seq_new:N} if necessary, then applies \texttt{\seq_(g)clear:N} to leave the \textit{(sequence)} empty.

\texttt{\seq_set_eq:NN \langle sequence\rangle_1 \langle sequence\rangle_2} \hspace{2cm} \texttt{\seq_set_eq:\ (cN|Nc|cc) \langle sequence\rangle_1 \langle sequence\rangle_2} \hspace{2cm} \texttt{\seq_gset_eq:NN \langle sequence\rangle_1 \langle sequence\rangle_2} \hspace{2cm} \texttt{\seq_gset_eq:\ (cN|Nc|cc) \langle sequence\rangle_1 \langle sequence\rangle_2}

Sets the content of \textit{(sequence\rangle_1} equal to that of \textit{(sequence\rangle_2}.

\texttt{\seq_set_from_clist:NN \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_set_from_clist:(cN|Nc|cc) \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_set_from_clist:Nn \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_set_from_clist:cn \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_gset_from_clist:NN \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_gset_from_clist:(cN|Nc|cc) \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_gset_from_clist:Nn \langle sequence\rangle \langle comma list\rangle} \hspace{2cm} \texttt{\seq_gset_from_clist:cn \langle sequence\rangle \langle comma list\rangle}

Converts the data in the \textit{(comma list)} into a \textit{(sequence)}: the original \textit{(comma list)} is unchanged.
\seq_const_from_clist:Nn \seq_const_from_clist:cn

Creates a new constant \langle \text{seq var} \rangle or raises an error if the name is already taken. The \langle \text{seq var} \rangle is set globally to contain the items in the \langle \text{comma list} \rangle.

\seq_set_split:Nnn \seq_set_split:NnV \seq_set_split:NnV

Splits the \langle \text{token list} \rangle into \langle \text{items} \rangle separated by \langle \text{delimiter} \rangle, and assigns the result to the \langle \text{sequence} \rangle. Spaces on both sides of each \langle \text{item} \rangle are ignored, then one set of outer braces is removed (if any): this space trimming behaviour is identical to that of \text{l3clist} functions. Empty \langle \text{items} \rangle are preserved by \seq_set_split:Nnn, and can be removed afterwards using \seq_remove_all:Nn \langle \text{sequence} \rangle {\langle \rangle}. The \langle \text{delimiter} \rangle may not contain \{, \} or \# (assuming \text{TeX}'s normal category code régime). If the \langle \text{delimiter} \rangle is empty, the \langle \text{token list} \rangle is split into \langle \text{items} \rangle as a \langle \text{token list} \rangle.

\seq_concat:NNN \seq_concat:ccc \seq_gconcat:NNN \seq_gconcat:ccc

Concatenates the content of \langle \text{sequence}_2 \rangle and \langle \text{sequence}_3 \rangle together and saves the result in \langle \text{sequence}_1 \rangle. The items in \langle \text{sequence}_2 \rangle are placed at the left side of the new sequence.

\seq_if_exist_p:N \seq_if_exist:NTF \seq_if_exist:cn \seq_if_exist:cv
\seq_if_exist_p:c \seq_if_exist:NTF c \seq_if_exist:cn c \seq_if_exist:cv c

Tests whether the \langle \text{sequence} \rangle is currently defined. This does not check that the \langle \text{sequence} \rangle really is a sequence variable.

2 Appending data to sequences

\seq_put_left:Nn \seq_put_left:NnV \seq_gput_left:Nn \seq_gput_left:NV \seq_gput_left:Nv \seq_gput_left:NVv \seq_gput_left:Nx \seq_gput_left:NVx \seq_gput_left:cn \seq_gput_left:cV \seq_gput_left:cv \seq_gput_left:cv \seq_gput_left:co \seq_gput_left:cx \seq_gput_left:cx
\seq_put_right:Nn \seq_put_right:NnV \seq_gput_right:Nn \seq_gput_right:NV \seq_gput_right:Nv \seq_gput_right:NVv \seq_gput_right:Nx \seq_gput_right:NVx \seq_gput_right:cn \seq_gput_right:cV \seq_gput_right:cv \seq_gput_right:cv \seq_gput_right:co \seq_gput_right:cx \seq_gput_right:cx

Appends the \langle \text{item} \rangle to the left of the \langle \text{sequence} \rangle.

3 Recovering items from sequences

Items can be recovered from either the left or the right of sequences. For implementation reasons, the actions at the left of the sequence are faster than those acting on the right. These functions all assign the recovered material locally, \emph{i.e.} setting the \langle \text{token list variable} \rangle used with \text{l3_set:Nn} and \emph{never} \text{l3_gset:Nn}.
\seq_get_left:NN \seq_get_left:cN
Stores the left-most item from a \seq \langle token list variable \rangle without removing it from the \seq. The \langle token list variable \rangle is assigned locally. If \seq is empty the \langle token list variable \rangle is set to the special marker \q_no_value.

\seq_get_right:NN \seq_get_right:cN
Stores the right-most item from a \seq \langle token list variable \rangle without removing it from the \seq. The \langle token list variable \rangle is assigned locally. If \seq is empty the \langle token list variable \rangle is set to the special marker \q_no_value.

\seq_pop_left:NN \seq_pop_left:cN
Pops the left-most item from a \seq \langle token list variable \rangle, i.e. removes the item from the sequence and stores it in the \langle token list variable \rangle. Both of the variables are assigned locally. If \seq is empty the \langle token list variable \rangle is set to the special marker \q_no_value.

\seq_gpop_left:NN \seq_gpop_left:cN
Pops the left-most item from a \seq \langle token list variable \rangle, i.e. removes the item from the sequence and stores it in the \langle token list variable \rangle. The \seq is modified globally, while the assignment of the \langle token list variable \rangle is local. If \seq is empty the \langle token list variable \rangle is set to the special marker \q_no_value.

\seq_pop_right:NN \seq_pop_right:cN
Pops the right-most item from a \seq \langle token list variable \rangle, i.e. removes the item from the sequence and stores it in the \langle token list variable \rangle. Both of the variables are assigned locally. If \seq is empty the \langle token list variable \rangle is set to the special marker \q_no_value.

\seq_gpop_right:NN \seq_gpop_right:cN
Pops the right-most item from a \seq \langle token list variable \rangle, i.e. removes the item from the sequence and stores it in the \langle token list variable \rangle. The \seq is modified globally, while the assignment of the \langle token list variable \rangle is local. If \seq is empty the \langle token list variable \rangle is set to the special marker \q_no_value.

\seq_item:Nn \seq_item:cn
Indexing items in the \seq from 1 at the top (left), this function evaluates the \langle integer expression \rangle and leaves the appropriate item from the sequence in the input stream. If the \langle integer expression \rangle is negative, indexing occurs from the bottom (right) of the sequence. If the \langle integer expression \rangle is larger than the number of items in the \seq (as calculated by \seq_count:N) then the function expands to nothing.

\mathrm{\TeXhackers \ note:} The result is returned within the \unexpanded primitive \exp_not:n, which means that the \langle item \rangle does not expand further when appearing in an \texttt{x}-type argument expansion.
\seq_rand_item:N \seq_rand_item:C *
Rev: 2016-12-06

\seq_rand_item:N \seq var
Selects a pseudo-random item of the \langle sequence \rangle. If the \langle sequence \rangle is empty the result is empty. This is not available in older versions of Xe\TeX.

\TeXhackers note: The result is returned within the \unexpanded primitive (\exp_not:n), which means that the \langle item \rangle does not expand further when appearing in an x-type argument expansion.

4 Recovering values from sequences with branching

The functions in this section combine tests for non-empty sequences with recovery of an item from the sequence. They offer increased readability and performance over separate testing and recovery phases.

\seq_get_left:NNTF \seq_get_left:NN \seq_get_left:cN *

If the \langle sequence \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle sequence \rangle is non-empty, stores the left-most item from the \langle sequence \rangle in the \langle token list variable \rangle without removing it from the \langle sequence \rangle, then leaves the \langle true code \rangle in the input stream. The \langle token list variable \rangle is assigned locally.

\seq_get_right:NNTF \seq_get_right:NN \seq_get_right:cN *
Rev: 2012-05-19

If the \langle sequence \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle sequence \rangle is non-empty, stores the right-most item from the \langle sequence \rangle in the \langle token list variable \rangle without removing it from the \langle sequence \rangle, then leaves the \langle true code \rangle in the input stream. The \langle token list variable \rangle is assigned locally.

\seq_pop_left:NNTF \seq_pop_left:NN \seq_pop_left:cN *

If the \langle sequence \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle sequence \rangle is non-empty, pops the left-most item from the \langle sequence \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle sequence \rangle, then leaves the \langle true code \rangle in the input stream. Both the \langle sequence \rangle and the \langle token list variable \rangle are assigned locally.

\seq_gpop_left:NNTF \seq_gpop_left:NN \seq_gpop_left:cN *

If the \langle sequence \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle sequence \rangle is non-empty, pops the left-most item from the \langle sequence \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle sequence \rangle, then leaves the \langle true code \rangle in the input stream. The \langle sequence \rangle is modified globally, while the \langle token list variable \rangle is assigned locally.
5 Modifying sequences

While sequences are normally used as ordered lists, it may be necessary to modify the content. The functions here may be used to update sequences, while retaining the order of the unaffected entries.

\seq_remove_duplicates:N \seq_remove_duplicates:c \seq_gremove_duplicates:N \seq_gremove_duplicates:c

\seq_remove_all:Nn \seq_remove_all:cn \seq_gremove_all:Nn \seq_gremove_all:cn

\seq_reverse:N \seq_reverse:c \seq_greverse:N \seq_greverse:c

\seq_sort:Nn \seq_sort:cn \seq_gsort:Nn \seq_gsort:cn

If the \langle sequence \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle sequence \rangle is non-empty, pops the right-most item from the \langle sequence \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle sequence \rangle, then leaves the \langle true code \rangle in the input stream. Both the \langle sequence \rangle and the \langle token list variable \rangle are assigned locally.

If the \langle sequence \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle sequence \rangle is non-empty, pops the right-most item from the \langle sequence \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle sequence \rangle, then leaves the \langle true code \rangle in the input stream. The \langle sequence \rangle is modified globally, while the \langle token list variable \rangle is assigned locally.

TEXhackers note: This function iterates through every item in the \langle sequence \rangle and does a comparison with the \langle items \rangle already checked. It is therefore relatively slow with large sequences.

Removes duplicate items from the \langle sequence \rangle, leaving the left most copy of each item in the \langle sequence \rangle. The \langle item \rangle comparison takes place on a token basis, as for \tl_if_eq:nnTF.

Removes every occurrence of \langle item \rangle from the \langle sequence \rangle. The \langle item \rangle comparison takes place on a token basis, as for \tl_if_eq:nnTF.

Reverses the order of the items stored in the \langle sequence \rangle.

Sorts the items in the \langle sequence \rangle according to the \langle comparison code \rangle, and assigns the result to \langle sequence \rangle. The details of sorting comparison are described in Section 1.

79
\seq_shuffle:N \seq_shuffle:c \seq_gshuffle:N \seq_gshuffle:c

\seq_shuffle:N \seq_shuffle:c

Sets the \textit{seq var} to the result of placing the items of the \textit{seq var} in a random order. Each item is (roughly) as likely to end up in any given position.

\textbf{\LaTeX{} hackers note:} For sequences with more than 13 items or so, only a small proportion of all possible permutations can be reached, because the random seed \sysrandseed: only has 28-bits. The use of \texttt{\toks} internally means that sequences with more than 32767 or 65535 items (depending on the engine) cannot be shuffled.

6 Sequence conditionals

\seq_if_empty_p:N \seq_if_empty_p:c \seq_if_empty:N \seq_if_empty:c

\seq_if_empty_p:N \seq_if_empty_p:c

Tests if the \textit{sequence} is empty (containing no items).

\seq_if_in:NnTF \seq_if_in:Nn \seq_if_in:NTF \seq_if_in:NTF \seq_if_in:NV \seq_if_in:Nv \seq_if_in:Nx \seq_if_in:cn \seq_if_in:cV \seq_if_in:cv \seq_if_in:co \seq_if_in:cx

Tests if the \textit{item} is present in the \textit{sequence}.

7 Mapping to sequences

All mappings are done at the current group level, \textit{i.e.} any local assignments made by the \textit{function} or \textit{code} discussed below remain in effect after the loop.

\seq_map_function:NN \seq_map_function:CN \seq_map_function:NN \seq_map_function:CN

\seq_map_function:NN \seq_map_function:CN

Applies \textit{function} to every \textit{item} stored in the \textit{sequence}. The \textit{function} will receive one argument for each iteration. The \textit{items} are returned from left to right. To pass further arguments to the \textit{function}, see \seqmaptokens:Nn. The function \seqmapinline:Nn is faster than \seqmapfunction:NN for sequences with more than about 10 items.

\seq_map_inl:Nn \seq_map_inl:cn

\seq_map_inl:Nn \seq_map_inl:cn

Applies \textit{inline function} to every \textit{item} stored within the \textit{sequence}. The \textit{inline function} should consist of code which will receive the \textit{item} as \#1. The \textit{items} are returned from left to right.
Analogue of \seq_map_function:NN which maps several tokens instead of a single function. The ⟨code⟩ receives each item in the ⟨sequence⟩ as two trailing brace groups. For instance,

\begin{verbatim}
\seq_map_tokens:Nn \l_my_seq { \prg_replicate:nn { 2 } }
\end{verbatim}

expands to twice each item in the ⟨sequence⟩: for each item in \l_my_seq the function \prg_replicate:nn receives 2 and ⟨item⟩ as its two arguments. The function \seq_map_inline:Nn is typically faster but is not expandable.

Stores each ⟨item⟩ of the ⟨sequence⟩ in turn in the (token list) ⟨variable⟩ and applies the ⟨code⟩. The ⟨code⟩ will usually make use of the ⟨variable⟩, but this is not enforced. The assignments to the ⟨variable⟩ are local. Its value after the loop is the last ⟨item⟩ in the ⟨sequence⟩, or its original value if the ⟨sequence⟩ is empty. The ⟨items⟩ are returned from left to right.

Used to terminate a \seq_map_... function before all entries in the ⟨sequence⟩ have been processed. This normally takes place within a conditional statement, for example

\begin{verbatim}
\seq_map_inline:Nn \l_my_seq
{ \str_if_eq:nnTF { #1 } { bingo } { \seq_map_break: } { % Do something useful }
}
\end{verbatim}

Use outside of a \seq_map_... scenario leads to low level \TeX{} errors.

\TeXhackers note: When the mapping is broken, additional tokens may be inserted before further items are taken from the input stream. This depends on the design of the mapping function.
\seq_map_break:n \{\langle\text{code}\rangle\}

Used to terminate a \seq_map... function before all entries in the \langle\text{sequence}\rangle have been processed, inserting the \langle\text{code}\rangle after the mapping has ended. This normally takes place within a conditional statement, for example

\seq_map_inline:Nn \l_my_seq
{\str_if_eq:nnTF { #1 } { bingo }{ \seq_map_break:n { <\text{code}> } }{\% Do something useful}}

Use outside of a \seq_map... scenario leads to low level \TeX{} errors.

\textbf{\TeX{}hackers note:} When the mapping is broken, additional tokens may be inserted before the \langle\text{code}\rangle is inserted into the input stream. This depends on the design of the mapping function.

\seq_count:N \langle\text{sequence}\rangle
Leaves the number of items in the \langle\text{sequence}\rangle in the input stream as an \langle\text{integer denotation}\rangle. The total number of items in a \langle\text{sequence}\rangle includes those which are empty and duplicates, \textit{i.e.} every item in a \langle\text{sequence}\rangle is unique.

8 Using the content of sequences directly

\seq_use:Nnnn \langle\text{seq var}\rangle \{\langle\text{separator between two}\rangle\}
\seq_use:Nnnn \langle\text{seq var}\rangle \{\langle\text{separator between more than two}\rangle\} \{\langle\text{separator between final two}\rangle\}

Places the contents of the \langle\text{seq var}\rangle in the input stream, with the appropriate \langle\text{separator}\rangle between the items. Namely, if the sequence has more than two items, the \langle\text{separator between more than two}\rangle is placed between each pair of items except the last, for which the \langle\text{separator between final two}\rangle is used. If the sequence has exactly two items, then they are placed in the input stream separated by the \langle\text{separator between two}\rangle. If the sequence has a single item, it is placed in the input stream, and an empty sequence produces no output. An error is raised if the variable does not exist or if it is invalid.

For example,

\seq_set_split:Nnn \l_tmpa_seq { | } { a | b | c | \{de\} | f }
\seq_use:Nnnn \l_tmpa_seq \{ ,~ and~\} \{ ,~\} \{ ,~and~\}

inserts “a, b, c, de, and f” in the input stream. The first separator argument is not used in this case because the sequence has more than 2 items.

\textbf{\TeX{}hackers note:} The result is returned within the \texttt{\textbackslash exp\textunderscore not:n}, which means that the \langle\text{items}\rangle do not expand further when appearing in an \texttt{x}-type argument expansion.
\seq_use:Nn \seq_use:cn *

Places the contents of the \seq_var{} in the input stream, with the \seq_separator{} between the items. If the sequence has a single item, it is placed in the input stream with no \seq_separator{}, and an empty sequence produces no output. An error is raised if the variable does not exist or if it is invalid.

For example,
\seq_set_split:Nnn \l_tmpa_seq { | } { a | b | c | \{de\} | f }
\seq_use:Nn \l_tmpa_seq { ~and~ }

inserts “a and b and c and de and f” in the input stream.

\TeX\_hackers\ note: The result is returned within the \unexpanded primitive (\exp_not:n), which means that the \items{} do not expand further when appearing in an \x\_type argument expansion.

9 Sequences as stacks

Sequences can be used as stacks, where data is pushed to and popped from the top of the sequence. (The left of a sequence is the top, for performance reasons.) The stack functions for sequences are not intended to be mixed with the general ordered data functions detailed in the previous section: a sequence should either be used as an ordered data type or as a stack, but not in both ways.

\seq_get:NN \seq_get:cN

Reads the top item from a \sequence{} into the \token_list_variable{} without removing it from the \sequence{}. The \token_list_variable{} is assigned locally. If \sequence{} is empty the \token_list_variable{} is set to the special marker \q_no_value.

\seq_pop:NN \seq_pop:cN

Pops the top item from a \sequence{} into the \token_list_variable{}. Both of the variables are assigned locally. If \sequence{} is empty the \token_list_variable{} is set to the special marker \q_no_value.

\seq_gpop:NN \seq_gpop:cN

Pops the top item from a \sequence{} into the \token_list_variable{}. The \sequence{} is modified globally, while the \token_list_variable{} is assigned locally. If \sequence{} is empty the \token_list_variable{} is set to the special marker \q_no_value.

\seq_get:NNTF \seq_get:cNTF

If the \sequence{} is empty, leaves the \false_code{} in the input stream. The value of the \token_list_variable{} is not defined in this case and should not be relied upon. If the \sequence{} is non-empty, stores the top item from a \sequence{} in the \token_list_variable{} without removing it from the \sequence{}. The \token_list_variable{} is assigned locally.
\seq_pop:NNTF \seq_pop:cNNTF

New: 2012-05-14
Updated: 2012-05-19

\seq_pop:NNTF \seq_pop:cNNTF

New: 2012-05-14
Updated: 2012-05-19

\seq_gpop:NNTF \seq_gpop:cNNTF

New: 2012-05-14
Updated: 2012-05-19

\seq_gpop:NN \seq_gpop:NNN

\seq_gpush:Nn \seq_gpush:NN

\seq_push:Nn \seq_push:Nn

\seq_push:Nn \seq_push:Nn

\seq_push:Nn \seq_push:NN

\seq_push:Nn \seq_push:NN

\seq_push:Nn \seq_push:Nn

Add the \{item\} to the top of the \langle sequence \rangle.

10 Sequences as sets

Sequences can also be used as sets, such that all of their items are distinct. Usage of sequences as sets is not currently widespread, hence no specific set function is provided. Instead, it is explained here how common set operations can be performed by combining several functions described in earlier sections. When using sequences to implement sets, one should be careful not to rely on the order of items in the sequence representing the set.

Sets should not contain several occurrences of a given item. To make sure that a \langle sequence variable \rangle only has distinct items, use \seq_remove_duplicates:N \langle sequence variable \rangle. This function is relatively slow, and to avoid performance issues one should only use it when necessary.

Some operations on a set \langle seq var \rangle are straightforward. For instance, \seq_count:N \langle seq var \rangle expands to the number of items, while \seq_if_in:NnTF \langle seq var \rangle \{\{item\}\} tests if the \{item\} is in the set.

Adding an \{item\} to a set \langle seq var \rangle can be done by appending it to the \langle seq var \rangle if it is not already in the \langle seq var \rangle:

\seq_if_in:NnF \langle seq var \rangle \{\{item\}\}
{ \seq_put_right:Nn \langle seq var \rangle \{\{item\}\} }

Removing an \{item\} from a set \langle seq var \rangle can be done using \seq_remove_all:Nn,

\seq_remove_all:Nn \langle seq var \rangle \{\{item\}\}

The intersection of two sets \langle seq var_1 \rangle and \langle seq var_2 \rangle can be stored into \langle seq var_3 \rangle by collecting items of \langle seq var_1 \rangle which are in \langle seq var_2 \rangle.
\seq_clear:N \l__{pkg}_internal_seq
\seq_map_inline:Nn \l__{pkg}_internal_seq {#1}
\seq_put_right:Nn \l__{pkg}_internal_seq {#1} }
}

The code as written here only works if \l__{seq} is different from the other two sequence variables. To cover all cases, items should first be collected in a sequence \l__{pkg}_internal_seq, then \l__{seq} should be set equal to this internal sequence. The same remark applies to other set functions.

The union of two sets \l__{seq} can be stored into \l__{seq} through
\seq_concat:NNN \l__{seq} \l__{seq} \l__{seq}
\seq_remove_duplicates:N \l__{seq}

or by adding items to (a copy of) \l__{seq} one by one
\seq_set_eq:NN \l__{seq} \l__{seq}
\seq_map_inline:Nn \l__{seq} { \seq_remove_all:Nn \l__{seq} {#1} }

The second approach is faster than the first when the \l__{seq} is short compared to \l__{seq}.

The difference of two sets \l__{seq} and \l__{seq} can be stored into \l__{seq} by removing items of the \l__{seq} from (a copy of) the \l__{seq} one by one.
\seq_set_eq:NN \l__{seq} \l__{seq}
\seq_map_inline:Nn \l__{seq} { \seq_remove_all:Nn \l__{seq} {#1} }

The symmetric difference of two sets \l__{seq} and \l__{seq} can be stored into \l__{seq} by computing the difference between \l__{seq} and \l__{seq} and storing the result as \l__{pkg}_internal_seq, then the difference between \l__{seq} and \l__{seq}, and finally concatenating the two differences to get the symmetric differences.
\seq_set_eq:NN \l__{seq} \l__{pkg}_internal_seq \l__{seq}
\seq_map_inline:Nn \l__{seq} { \seq_remove_all:Nn \l__{seq} {#1} }

\seq_set_eq:NN \l__{seq} \l__{seq}
\seq_map_inline:Nn \l__{seq} { \seq_remove_all:Nn \l__{seq} {#1} }
\seq_map_inline:Nn \l__{seq} { \seq_remove_all:Nn \l__{seq} {#1} }
\seq_set_eq:NN \l__{seq} \l__{seq} \l__{pkg}_internal_seq

11 Constant and scratch sequences

\c_empty_seq Constant that is always empty.

\seq: 2012-07-02

85
Scratch sequences for local assignment. These are never used by the kernel code, and so are safe for use with any \texttt{TEX3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\l_tmpa_seq
\l_tmpb_seq
New: 2012-04-26

Scratch sequences for global assignment. These are never used by the kernel code, and so are safe for use with any \texttt{TEX3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_seq
\g_tmpb_seq
New: 2012-04-26

12 Viewing sequences

\seq_show:N \seq_show:c
Displays the entries in the \texttt{sequence} in the terminal.

\seq_log:N \seq_log:c
Writes the entries in the \texttt{sequence} in the log file.
Part XI
The l3int package
Integers

Calculation and comparison of integer values can be carried out using literal numbers, int
registers, constants and integers stored in token list variables. The standard operators
+, -, / and * and parentheses can be used within such expressions to carry arithmetic
operations. This module carries out these functions on integer expressions (“intexpr”).
Integer expressions

\int_eval:n \{ \langle integer expression \rangle \}

Evaluates the \langle integer expression \rangle and leaves the result in the input stream as an integer denotation: for positive results an explicit sequence of decimal digits not starting with 0, for negative results \(-\) followed by such a sequence, and 0 for zero. The \langle integer expression \rangle should consist, after expansion, of \(+\), \(-\), \(*\), \(/\), \(\langle\), \(\rangle\) and of course integer operands. The result is calculated by applying standard mathematical rules with the following peculiarities:

- \(/\) denotes division rounded to the closest integer with ties rounded away from zero;
- there is an error and the overall expression evaluates to zero whenever the absolute value of any intermediate result exceeds \(2^{31} - 1\), except in the case of scaling operations \(a*b/c\), for which \(a*b\) may be arbitrarily large;
- parentheses may not appear after unary \(+\) or \(-\), namely placing \(+\langle\) or \(-\langle\ at the start of an expression or after \(+\), \(-\), \(*\), \(/\) or \(\langle\ leads to an error.

Each integer operand can be either an integer variable (with no need for \int_use:N) or an integer denotation. For example both

\int_eval:n \{ 5 + 4 * 3 - ( 3 + 4 * 5 ) \}

and

\tl_new:N \l_my_tl
\tl_set:Nn \l_my_tl { 5 }
\int_new:N \l_my_int
\int_set:Nn \l_my_int { 4 }
\int_eval:n \{ \l_my_tl + \l_my_int * 3 - ( 3 + 4 * 5 ) \}

evaluate to \(-6\) because \(\l_my_tl\) expands to the integer denotation \(5\). As the \langle integer expression \rangle is fully expanded from left to right during evaluation, fully expandable and restricted-expandable functions can both be used, and \exp_not:n and its variants have no effect while \exp_not:N may incorrectly interrupt the expression.

\TeXhackers note: Exactly two expansions are needed to evaluate \int_eval:n. The result is not an \langle internal integer \rangle, and therefore requires suitable termination if used in a \TeXX-style integer assignment.

As all \TeXX integers, integer operands can also be dimension or skip variables, converted to integers in \sp, or octal numbers given as \' followed by digits other than 8 and 9, or hexadecimal numbers given as " followed by digits or upper case letters from A to F, or the character code of some character or one-character control sequence, given as \'\langle char\rangle.
\int_eval:w \int_eval:w \langle \text{integer expression} \rangle

Evaluates the \langle \text{integer expression} \rangle as described for \int_eval:n. The end of the expression is the first token encountered that cannot form part of such an expression. If that token is \scan_stop: it is removed, otherwise not. Spaces do not terminate the expression. However, spaces terminate explicit integers, and this may terminate the expression: for instance, \int_eval:w \text{1}_4+\text{1}_9\text{9} expands to \text{29} since the digit \text{9} is not part of the expression.

\int_sign:n \{\langle \text{integer expression} \rangle\}

New: 2018-11-03

Evaluates the \langle \text{integer expression} \rangle then leaves 1 or 0 or \text{−1} in the input stream according to the sign of the result.

\int_abs:n \{\langle \text{integer expression} \rangle\}

Updated: 2012-09-26

Evaluates the \langle \text{integer expression} \rangle as described for \int_eval:n and leaves the absolute value of the result in the input stream as an \langle \text{integer denotation} \rangle after two expansions.

\int_div_round:nn \{\langle \text{integer expression} \rangle\} \{\langle \text{integer expression} \rangle\}

Updated: 2012-09-26

Evaluates the two \langle \text{integer expressions} \rangle as described earlier, then divides the first value by the second, and rounds the result to the closest integer. Ties are rounded away from zero. Note that this is identical to using / directly in an \langle \text{integer expression} \rangle. The result is left in the input stream as an \langle \text{integer denotation} \rangle after two expansions.

\int_div_truncate:nn \{\langle \text{integer expression} \rangle\} \{\langle \text{integer expression} \rangle\}

Updated: 2012-02-09

Evaluates the two \langle \text{integer expressions} \rangle as described earlier, then divides the first value by the second, and rounds the result towards zero. Note that division using / rounds to the closest integer instead. The result is left in the input stream as an \langle \text{integer denotation} \rangle after two expansions.

\int_max:nn \{\langle \text{integer expression} \rangle\} \{\langle \text{integer expression} \rangle\}

\int_min:nn \{\langle \text{integer expression} \rangle\} \{\langle \text{integer expression} \rangle\}

Updated: 2012-09-26

Evaluates the \langle \text{integer expressions} \rangle as described for \int_eval:n and leaves either the larger or smaller value in the input stream as an \langle \text{integer denotation} \rangle after two expansions.

\int_mod:nn \{\langle \text{integer expression} \rangle\} \{\langle \text{integer expression} \rangle\}

Updated: 2012-09-26

Evaluates the two \langle \text{integer expressions} \rangle as described earlier, then calculates the integer remainder of dividing the first expression by the second. This is obtained by subtracting \int_div_truncate:nn \{\langle \text{integer expression} \rangle\} \{\langle \text{integer expression} \rangle\} times \langle \text{integer expression} \rangle from \langle \text{integer expression} \rangle. Thus, the result has the same sign as \langle \text{integer expression} \rangle and its absolute value is strictly less than that of \langle \text{integer expression} \rangle. The result is left in the input stream as an \langle \text{integer denotation} \rangle after two expansions.

2 Creating and initialising integers

\int_new:N \int_new:N \langle \text{integer} \rangle

Creates a new \langle \text{integer} \rangle or raises an error if the name is already taken. The declaration is global. The \langle \text{integer} \rangle is initially equal to 0.
\int_const:Nn
\int_const:cn

Updated: 2011-10-22

\int_zero:N
\int_zero:c
\int_gzero:N
\int_gzero:c

\int_zero_new:N
\int_zero_new:c
\int_gzero_new:N
\int_gzero_new:c

New: 2011-12-13

\int_set_eq:NN
\int_set_eq:(cN|Nc|cc)
\int_gset_eq:NN
\int_gset_eq:(cN|Nc|cc)

\int_if_exist_p:N
\int_if_exist_p:c
\int_if_exist:NTF
\int_if_exist:cTF

New: 2012-03-03

3 Setting and incrementing integers

\int_add:Nn
\int_add:cn
\int_gadd:Nn
\int_gadd:cn

Updated: 2011-10-22

\int_decr:N
\int_decr:c
\int_gdecr:N
\int_gdecr:c

\int_incr:N
\int_incr:c
\int_gincr:N
\int_gincr:c

\int_const:Nn \{\langle integer expression\rangle\}

Creates a new constant \langle integer\rangle or raises an error if the name is already taken. The value of the \langle integer\rangle is set globally to the \langle integer expression\rangle.

\int_zero:N \{\langle integer\rangle\}

Sets \langle integer\rangle to 0.

\int_zero_new:N \{\langle integer\rangle\}

Ensures that the \langle integer\rangle exists globally by applying \int_new:N if necessary, then applies \int_)zero:N to leave the \langle integer\rangle set to zero.

\int_set_eq:NN \{\langle integer1\rangle\} \{\langle integer2\rangle\}

Sets the content of \langle integer1\rangle equal to that of \langle integer2\rangle.

\int_if_exist_p:N \langle \langle int\rangle \rangle
\int_if_exist:NTF \langle \langle int\rangle \rangle \{\langle true\ code\rangle\} \{\langle false\ code\rangle\}

Tests whether the \langle int\rangle is currently defined. This does not check that the \langle int\rangle really is an integer variable.

\int_add:Nn \{\langle integer expression\rangle\}

Adds the result of the \langle integer expression\rangle to the current content of the \langle integer\rangle.

\int_decr:N \{\langle integer\rangle\}

Decreases the value stored in \langle integer\rangle by 1.

\int_incr:N \{\langle integer\rangle\}

Increases the value stored in \langle integer\rangle by 1.
\int_set:Nn \int_set:cn \int_gset:Nn \int_gset:cn

Sets \texttt{integer} to the value of \texttt{integer expression}, which must evaluate to an integer (as described for \texttt{int_eval:n}).

\int_sub:Nn \int_sub:cn \int_gsub:Nn \int_gsub:cn

Updated: 2011-10-22

\int_set:Nn \int_set:cn \int_gset:Nn \int_gset:cn

Subtracts the result of the \texttt{integer expression} from the current content of the \texttt{integer}.

\int_use:N \int_use:c

Updated: 2011-10-22

\int_use:N \int_use:c

4 Using integers

\int_use:N \int_use:cn

Recovers the content of an \texttt{integer} and places it directly in the input stream. An error is raised if the variable does not exist or if it is invalid. Can be omitted in places where an \texttt{integer} is required (such as in the first and third arguments of \texttt{int_compare:nNnTF}).

\textbf{\TeX}hackers note: \texttt{int_use:N} is the \TeX primitive \texttt{the}: this is one of several \LaTeX names for this primitive.

5 Integer expression conditionals

\int_compare_p:nNn  \int_compare:nnTF

\int_compare_p:nNn  \int_compare:nnTF

\int_compare_p:nNn  \int_compare:nnTF

This function first evaluates each of the \texttt{integer expressions} as described for \texttt{int_eval:n}. The two results are then compared using the \texttt{relation}:

- Equal $=$
- Greater than $>$
- Less than $<$

This function is less flexible than \texttt{int_compare:nTF} but around 5 times faster.
This function evaluates the \textit{integer expressions} as described for \texttt{\textbackslash int\_eval:n} and compares consecutive result using the corresponding \textit{relation}, namely it compares \texttt{\intexpr_1} and \texttt{\intexpr_2} using the \texttt{\relation_1}, then \texttt{\intexpr_2} and \texttt{\intexpr_3} using the \texttt{\relation_2}, until finally comparing \texttt{\intexpr_N} and \texttt{\intexpr_{N+1}} using the \texttt{\relation_N}. The test yields \texttt{true} if all comparisons are \texttt{true}. Each \textit{integer expression} is evaluated only once, and the evaluation is lazy, in the sense that if one comparison is \texttt{false}, then no other \textit{integer expression} is evaluated and no other comparison is performed. The \textit{relations} can be any of the following:

\begin{itemize}
  \item Equal \quad = \text{ or } ==
  \item Greater than or equal to \quad >=
  \item Greater than \quad >
  \item Less than or equal to \quad <=
  \item Less than \quad <
  \item Not equal \quad !=
\end{itemize}

This function is more flexible than \texttt{\textbackslash int\_compare:nNnTF} but around 5 times slower.
This function evaluates the \textit{(test integer expression)} and compares this in turn to each of the \textit{(integer expression cases)}. If the two are equal then the associated \textit{(code)} is left in the input stream and other cases are discarded. If any of the cases are matched, the \textit{(true code)} is also inserted into the input stream (after the code for the appropriate case), while if none match then the \textit{(false code)} is inserted. The function \texttt{\int_case:nn}, which does nothing if there is no match, is also available. For example

\begin{verbatim}
\int_case:nnF 
{ 2 \ast 5 }
{ 
    { 5 } \{ Small \}
    { 4 + 6 } \{ Medium \}
    { -2 \ast 10 } \{ Negative \}
}
{ No idea! }
\end{verbatim}

leaves “Medium” in the input stream.

\begin{verbatim}
\int_if_even_p:n \int_if_odd_p:n
\int_if_odd:nTF \int_if_even:nTF 
\int_if_odd:n \int_if_even:n
\end{verbatim}

This function first evaluates the \textit{(integer expression)} as described for \texttt{\inteval:n}. It then evaluates if this is odd or even, as appropriate.

\section{Integer expression loops}

\begin{verbatim}
\int_do_until:nNnn 
\int_do_until:nNnn 
\int_do_while:nNnn 
\int_do_while:nNnn
\end{verbatim}

Places the \textit{(code)} in the input stream for \LaTeX{} to process, and then evaluates the relationship between the two \textit{(integer expressions)} as described for \texttt{\intcompare:nNnTF}. If the test is \texttt{false} then the \textit{(code)} is inserted into the input stream again and a loop occurs until the \textit{(relation)} is \texttt{true}.

Places the \textit{(code)} in the input stream for \LaTeX{} to process, and then evaluates the relationship between the two \textit{(integer expressions)} as described for \texttt{\intcompare:nNnTF}. If the test is \texttt{true} then the \textit{(code)} is inserted into the input stream again and a loop occurs until the \textit{(relation)} is \texttt{false}. 

93
\texttt{\textbackslash int\_until\_do:nNnn} ☆ \texttt{\textbackslash int\_until\_do:nNnn \{\text{intexpr}_1\} \{\text{relation}\} \{\text{intexpr}_2\} \{\text{code}\}}

Evaluates the relationship between the two \textit{integer expressions} as described for \texttt{\textbackslash int\_compare:nNnTF}, and then places the \texttt{\text{code}} in the input stream if the \texttt{\text{relation}} is \texttt{false}. After the \texttt{\text{code}} has been processed by \TeX the test is repeated, and a loop occurs until the test is \texttt{true}.

\texttt{\textbackslash int\_while\_do:nNnn} ☆ \texttt{\textbackslash int\_while\_do:nNnn \{\text{intexpr}_1\} \{\text{relation}\} \{\text{intexpr}_2\} \{\text{code}\}}

Evaluates the relationship between the two \textit{integer expressions} as described for \texttt{\textbackslash int\_compare:nNnTF}, and then places the \texttt{\text{code}} in the input stream if the \texttt{\text{relation}} is \texttt{true}. After the \texttt{\text{code}} has been processed by \TeX the test is repeated, and a loop occurs until the test is \texttt{false}.

\texttt{\textbackslash int\_do\_until:nn} ☆ \texttt{\textbackslash int\_do\_until:nn \{\text{integer relation}\} \{\text{code}\}}

Updated: 2013-01-13

Places the \texttt{\text{code}} in the input stream for \TeX to process, and then evaluates the \textit{integer relation} as described for \texttt{\textbackslash int\_compare:nNnTF}. If the test is \texttt{false} then the \texttt{\text{code}} is inserted into the input stream again and a loop occurs until the \texttt{\text{relation}} is \texttt{true}.

\texttt{\textbackslash int\_do\_while:nn} ☆ \texttt{\textbackslash int\_do\_while:nn \{\text{integer relation}\} \{\text{code}\}}

Updated: 2013-01-13

Places the \texttt{\text{code}} in the input stream for \TeX to process, and then evaluates the \textit{integer relation} as described for \texttt{\textbackslash int\_compare:nNnTF}. If the test is \texttt{true} then the \texttt{\text{code}} is inserted into the input stream again and a loop occurs until the \texttt{\text{relation}} is \texttt{false}.

\texttt{\textbackslash int\_until\_do:nn} ☆ \texttt{\textbackslash int\_until\_do:nn \{\text{integer relation}\} \{\text{code}\}}

Updated: 2013-01-13

Evaluates the \textit{integer relation} as described for \texttt{\textbackslash int\_compare:nNnTF}, and then places the \texttt{\text{code}} in the input stream if the \texttt{\text{relation}} is \texttt{false}. After the \texttt{\text{code}} has been processed by \TeX the test is repeated, and a loop occurs until the test is \texttt{true}.

\texttt{\textbackslash int\_while\_do:nn} ☆ \texttt{\textbackslash int\_while\_do:nn \{\text{integer relation}\} \{\text{code}\}}

Updated: 2013-01-13

Evaluates the \textit{integer relation} as described for \texttt{\textbackslash int\_compare:nNnTF}, and then places the \texttt{\text{code}} in the input stream if the \texttt{\text{relation}} is \texttt{true}. After the \texttt{\text{code}} has been processed by \TeX the test is repeated, and a loop occurs until the test is \texttt{false}.
7 Integer step functions

\texttt{int\_step\_function:nN} \quad \texttt{int\_step\_function:nnN} \quad \texttt{int\_step\_function:nnN}

This function first evaluates the \texttt{(initial value)}, \texttt{(step)} and \texttt{(final value)}, all of which should be integer expressions. The \texttt{(function)} is then placed in front of each \texttt{(value)} from the \texttt{(initial value)} to the \texttt{(final value)} in turn (using \texttt{(step)} between each \texttt{(value)}). The \texttt{(step)} must be non-zero. If the \texttt{(step)} is positive, the loop stops when the \texttt{(value)} becomes larger than the \texttt{(final value)}. If the \texttt{(step)} is negative, the loop stops when the \texttt{(value)} becomes smaller than the \texttt{(final value)}. The \texttt{(function)} should absorb one numerical argument. For example

\begin{verbatim}
\cs_set:Npn \my_func:n #1 { \[I saw #1\] \quad }
\int_step_function:nnnN { 1 } { 1 } { 5 } \my_func:n
\end{verbatim}

would print

\begin{verbatim}
[I saw 1] [I saw 2] [I saw 3] [I saw 4] [I saw 5]
\end{verbatim}

The functions \texttt{int\_step\_function:nN} and \texttt{int\_step\_function:nnN} both use a fixed \texttt{(step)} of 1, and in the case of \texttt{int\_step\_function:nN} the \texttt{(initial value)} is also fixed as 1. These functions are provided as simple short-cuts for code clarity.

\begin{verbatim}
\int_step_inline:nnn \my_func:n #1 { [I-saw-#1] \quad }
\end{verbatim}

This function first evaluates the \texttt{(initial value)}, \texttt{(step)} and \texttt{(final value)}, all of which should be integer expressions. Then for each \texttt{(value)} from the \texttt{(initial value)} to the \texttt{(final value)} in turn (using \texttt{(step)} between each \texttt{(value)}), the \texttt{(code)} is inserted into the input stream with \texttt{\#1} replaced by the current \texttt{(value)}. Thus the \texttt{(code)} should define a function of one argument (\texttt{\#1}).

The functions \texttt{int\_step\_inline:nn} and \texttt{int\_step\_inline:nnn} both use a fixed \texttt{(step)} of 1, and in the case of \texttt{int\_step\_inline:nn} the \texttt{(initial value)} is also fixed as 1. These functions are provided as simple short-cuts for code clarity.

\begin{verbatim}
\int_step_variable:nN \my_func:n #1 { [I-saw-#1] \quad }
\end{verbatim}

This function first evaluates the \texttt{(initial value)}, \texttt{(step)} and \texttt{(final value)}, all of which should be integer expressions. Then for each \texttt{(value)} from the \texttt{(initial value)} to the \texttt{(final value)} in turn (using \texttt{(step)} between each \texttt{(value)}), the \texttt{(code)} is inserted into the input stream, with the \texttt{(tl var)} defined as the current \texttt{(value)}. Thus the \texttt{(code)} should make use of the \texttt{(tl var)}.

The functions \texttt{int\_step\_variable:nN} and \texttt{int\_step\_variable:nnNn} both use a fixed \texttt{(step)} of 1, and in the case of \texttt{int\_step\_variable:nNn} the \texttt{(initial value)} is also fixed as 1. These functions are provided as simple short-cuts for code clarity.
8 Formatting integers

Integers can be placed into the output stream with formatting. These conversions apply to any integer expressions.

\int_to_arabic:n \{\langle integer expression\rangle\}

Places the value of the \langle integer expression\rangle in the input stream as digits, with category code 12 (other).

\int_to_alph:n \{\langle integer expression\rangle\}

Evaluates the \langle integer expression\rangle and converts the result into a series of letters, which are then left in the input stream. The conversion rule uses the 26 letters of the English alphabet, in order, adding letters when necessary to increase the total possible range of representable numbers. Thus

\int_to_alph:n \{ 1 \}

places a in the input stream,

\int_to_alph:n \{ 26 \}

is represented as z and

\int_to_alph:n \{ 27 \}

is converted to aa. For conversions using other alphabets, use \int_to_symbols:nnn to define an alphabet-specific function. The basic \int_to_alph:n and \int_to_Alph:n functions should not be modified. The resulting tokens are digits with category code 12 (other) and letters with category code 11 (letter).

\int_to_symbols:nnn \{\langle integer expression\rangle\} \{(total symbols)\} \{(value to symbol mapping)\}

This is the low-level function for conversion of an \langle integer expression\rangle into a symbolic form (often letters). The \langle total symbols\rangle available should be given as an integer expression. Values are actually converted to symbols according to the \langle value to symbol mapping\rangle. This should be given as \langle total symbols\rangle pairs of entries, a number and the appropriate symbol. Thus the \int_to_alph:n function is defined as

\cs_new:Npn \int_to_alph:n #1
{
  \int_to_symbols:nnn {#1} { 26 }
  \{ 1 \} \{ a \}
  \{ 2 \} \{ b \}
  \ldots
  \{ 26 \} \{ z \}
}
\int_to_bin:n \{\text{integer expression}\}
Calculates the value of the \textit{integer expression} and places the binary representation of the result in the input stream.

\int_to_hex:n \{\text{integer expression}\}
\int_to_Hex:n
Calculates the value of the \textit{integer expression} and places the hexadecimal (base 16) representation of the result in the input stream. Letters are used for digits beyond 9: lower case letters for \texttt{\int_to_hex:n} and upper case ones for \texttt{\int_to_Hex:n}. The resulting tokens are digits with category code 12 (other) and letters with category code 11 (letter).

\int_to_oct:n \{\text{integer expression}\}
Calculates the value of the \textit{integer expression} and places the octal (base 8) representation of the result in the input stream. The resulting tokens are digits with category code 12 (other) and letters with category code 11 (letter).

\int_to_base:nn \{\text{integer expression}\} \{\text{base}\}
\int_to_Base:nn
Calculates the value of the \textit{integer expression} and converts it into the appropriate representation in the \textit{base}; the later may be given as an integer expression. For bases greater than 10 the higher “digits” are represented by letters from the English alphabet: lower case letters for \texttt{\int_to_base:n} and upper case ones for \texttt{\int_to_Base:n}. The maximum \textit{base} value is 36. The resulting tokens are digits with category code 12 (other) and letters with category code 11 (letter).

\int_to_roman:n \{\text{integer expression}\}
\int_to_Roman:n
Places the value of the \textit{integer expression} in the input stream as Roman numerals, either lower case \texttt{\int_to_roman:n} or upper case \texttt{\int_to_Roman:n}. If the value is negative or zero, the output is empty. The Roman numerals are letters with category code 11 (letter). The letters used are \texttt{mdclxvi}, repeated as needed: the notation with bars (such as \texttt{\bar{v}} for 5000) is \textit{not} used. For instance \texttt{\int_to_roman:n \{ 8249 \}} expands to \texttt{mmmmmmmmccxlix}.

\textbf{\textsf{\LaTeX}hackers note:} This is a generic version of \texttt{\int_to_bin:n, etc.}

\int_from_alph:n \{\text{letters}\}
Converts the \textit{letters} into the integer (base 10) representation and leaves this in the input stream. The \textit{letters} are first converted to a string, with no expansion. Lower and upper case letters from the English alphabet may be used, with “a” equal to 1 through to “z” equal to 26. The function also accepts a leading sign, made of + and -. This is the inverse function of \texttt{\int_to_alph:n} and \texttt{\int_to_Alph:n}.  

9 Converting from other formats to integers
\texttt{\textbackslash int\_from\_bin:nn} \texttt{\{\langle\text{binary number}\rangle\}}

Converts the \textit{\langle binary number\rangle} into the integer (base 10) representation and leaves this in the input stream. The \textit{\langle binary number\rangle} is first converted to a string, with no expansion. The function accepts a leading sign, made of + and -, followed by binary digits. This is the inverse function of \texttt{\textbackslash int\_to\_bin:nn}.

\texttt{\textbackslash int\_from\_hex:nn} \texttt{\{\langle\text{hexadecimal number}\rangle\}}

Converts the \textit{\langle hexadecimal number\rangle} into the integer (base 10) representation and leaves this in the input stream. Digits greater than 9 may be represented in the \textit{\langle hexadecimal number\rangle} by upper or lower case letters. The \textit{\langle hexadecimal number\rangle} is first converted to a string, with no expansion. The function also accepts a leading sign, made of + and -. This is the inverse function of \texttt{\textbackslash int\_to\_hex:nn} and \texttt{\textbackslash int\_to\_Hex:nn}.

\texttt{\textbackslash int\_from\_oct:nn} \texttt{\{\langle\text{octal number}\rangle\}}

Converts the \textit{\langle octal number\rangle} into the integer (base 10) representation and leaves this in the input stream. The \textit{\langle octal number\rangle} is first converted to a string, with no expansion. The function accepts a leading sign, made of + and -, followed by octal digits. This is the inverse function of \texttt{\textbackslash int\_to\_oct:nn}.

\texttt{\textbackslash int\_from\_roman:nn} \texttt{\{\langle\text{roman numeral}\rangle\}}

Converts the \textit{\langle roman numeral\rangle} into the integer (base 10) representation and leaves this in the input stream. The \textit{\langle roman numeral\rangle} is first converted to a string, with no expansion. The \textit{\langle roman numeral\rangle} may be in upper or lower case; if the numeral contains characters besides mdclxvi or MDCLXVI then the resulting value is $-1$. This is the inverse function of \texttt{\textbackslash int\_to\_roman:nn} and \texttt{\textbackslash int\_to\_Roman:nn}.

\texttt{\textbackslash int\_from\_base:nn} \texttt{\{\langle\text{number}\rangle\} \{\langle\text{base}\rangle\}}

Converts the \textit{\langle number\rangle} expressed in \textit{\langle base\rangle} into the appropriate value in base 10. The \textit{\langle number\rangle} is first converted to a string, with no expansion. The \textit{\langle number\rangle} should consist of digits and letters (either lower or upper case), plus optionally a leading sign. The maximum \textit{\langle base\rangle} value is 36. This is the inverse function of \texttt{\textbackslash int\_to\_base:nn} and \texttt{\textbackslash int\_-\_to\_Base:nn}.

10 Random integers

\texttt{\textbackslash int\_rand:nn} \texttt{\{\langle\text{intexpr}\rangle\} \{\langle\text{intexpr}\rangle\}}

Evaluates the two \textit{\langle integer expressions\rangle} and produces a pseudo-random number between the two (with bounds included). This is not available in older versions of \texttt{Xe\TeX}.

\texttt{\textbackslash int\_rand:n} \texttt{\{\langle\text{intexpr}\rangle\}}

Evaluates the \textit{\langle integer expression\rangle} then produces a pseudo-random number between 1 and the \textit{\langle intexpr\rangle} (included). This is not available in older versions of \texttt{Xe\TeX}. 

98
11 Viewing integers

\int_show:N \langle integer \rangle
Displays the value of the \langle integer \rangle on the terminal.

\int_show:c \langle integer \rangle
Displays the result of evaluating the \langle integer expression \rangle on the terminal.

\int_log:N \langle integer \rangle
\int_log:c
\langle integer \rangle
Writes the value of the \langle integer \rangle in the log file.

\int_log:n \langle integer expression \rangle
\langle integer expression \rangle
Writes the result of evaluating the \langle integer expression \rangle in the log file.

12 Constant integers

\c_zero_int \c_one_int

The maximum value that can be stored as an integer.

\c_max_int
Maximum number of registers.

\c_max_char_int
Maximum character code completely supported by the engine.

13 Scratch integers

\l_tmpa_int \l_tmpb_int
Scratch integer for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_int \g_tmpb_int
Scratch integer for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
13.1 Direct number expansion

\verb|\int_value:w * \int_value:w ⟨integer⟩|
\verb|\int_value:w ⟨integer denotation⟩ ⟨optional space⟩|

Expands the following tokens until an \textit{⟨integer⟩} is formed, and leaves a normalized form
(no leading sign except for negative numbers, no leading digit 0 except for zero) in the
input stream as category code 12 (other) characters. The \textit{⟨integer⟩} can consist of any
number of signs (with intervening spaces) followed by
\begin{itemize}
  \item an integer variable (in fact, any \TeX{} register except \verb|\toks|)
  \item explicit digits (or by ‘\textit{⟨octal digits⟩}’ or “\textit{⟨hexadecimal digits⟩}’ or ‘\textit{⟨character⟩}).
\end{itemize}

In this last case expansion stops once a non-digit is found; if that is a space it is removed
as in \texttt{f}-expansion, and so \texttt{\exp_stop_f:} may be employed as an end marker. Note that
protected functions are expanded by this process.

This function requires exactly one expansion to produce a value, and so is suitable
for use in cases where a number is required “directly”. In general, \texttt{\int_eval:n} is the
preferred approach to generating numbers.

\TeX{}hackers note: This is the \TeX{} primitive \texttt{\number}.

14 Primitive conditionals

\verb|\if_int_compare:w * \if_int_compare:w ⟨integer₁⟩ ⟨relation⟩ ⟨integer₂⟩|
\verb|\true code⟩|
\verb|\else: |
\verb|\false code⟩|
\verb|\fi: |

Compare two integers using \textit{⟨relation⟩}, which must be one of =, < or > with category code
12. The \texttt{\else:} branch is optional.

\TeX{}hackers note: These are both names for the \TeX{} primitive \texttt{\ifnum}.

\begin{verbatim}
\if_case:w \or: \or: \else: \fi:
\end{verbatim}

Selects a case to execute based on the value of the \textit{⟨integer⟩}. The first case \textit{⟨case₀⟩)
}

may be a literal, a constant or an integer expression (\textit{e.g.} using \texttt{\int_eval:n}).

\TeX{}hackers note: These are the \TeX{} primitives \texttt{\ifcase} and \texttt{\or}.

100
\texttt{\textbackslash if\_int\_odd:w} *  
\texttt{\textbackslash if\_int\_odd:w} (tokens) (optional space) 
\texttt{\textbackslash if\_int\_odd:w} (true code) 
\texttt{\textbackslash else:} 
\texttt{\textbackslash else:} 
\texttt{\textbackslash true code} 
\texttt{\textbackslash fi:} 

Expands (tokens) until a non-numeric token or a space is found, and tests whether the resulting (integer) is odd. If so, (true code) is executed. The \texttt{\textbackslash else:} branch is optional.

\textbf{TeXhackers note:} This is the \TeX{} primitive \texttt{\textbackslash ifodd}. 
Part XII

The \l3flag package: Expandable flags

Flags are the only data-type that can be modified in expansion-only contexts. This module is meant mostly for kernel use: in almost all cases, booleans or integers should be preferred to flags because they are very significantly faster.

A flag can hold any non-negative value, which we call its ⟨height⟩. In expansion-only contexts, a flag can only be “raised”: this increases the ⟨height⟩ by 1. The ⟨height⟩ can also be queried expandably. However, decreasing it, or setting it to zero requires non-expandable assignments.

Flag variables are always local. They are referenced by a ⟨flag name⟩ such as \texttt{str\textunderscore missing}. The ⟨flag name⟩ is used as part of \texttt{\use: c} constructions hence is expanded at point of use. It must expand to character tokens only, with no spaces.

A typical use case of flags would be to keep track of whether an exceptional condition has occurred during expandable processing, and produce a meaningful (non-expandable) message after the end of the expandable processing. This is exemplified by \l3str\textunderscore convert, which for performance reasons performs conversions of individual characters expandably and for readability reasons produces a single error message describing incorrect inputs that were encountered.

Flags should not be used without carefully considering the fact that raising a flag takes a time and memory proportional to its height. Flags should not be used unless unavoidable.

1 Setting up flags

\begin{verbatim}
\flag_new:n \{⟨flag name⟩\}
\end{verbatim}

Creates a new flag with a name given by ⟨flag name⟩, or raises an error if the name is already taken. The ⟨flag name⟩ may not contain spaces. The declaration is global, but flags are always local variables. The ⟨flag⟩ initially has zero height.

\begin{verbatim}
\flag_clear:n \{⟨flag name⟩\}
\end{verbatim}

The ⟨flag⟩’s height is set to zero. The assignment is local.

\begin{verbatim}
\flag_clear_new:n \{⟨flag name⟩\}
\end{verbatim}

Ensures that the ⟨flag⟩ exists globally by applying \flag_new:n if necessary, then applies \flag_clear:n, setting the height to zero locally.

\begin{verbatim}
\flag_show:n \{⟨flag name⟩\}
\end{verbatim}

Displays the ⟨flag⟩’s height in the terminal.

\begin{verbatim}
\flag_log:n \{⟨flag name⟩\}
\end{verbatim}

Writes the ⟨flag⟩’s height to the log file.
2 Expandable flag commands

\flag_if_exist:n \{\textit{flag name}\} \\
\flag_if_exist:nTF

This function returns \texttt{true} if the \textit{flag name} references a flag that has been defined previously, and \texttt{false} otherwise.

\flag_if_raised:n \{\textit{flag name}\} \\
\flag_if_raised:nTF

This function returns \texttt{true} if the \textit{flag} has non-zero height, and \texttt{false} if the \textit{flag} has zero height.

\flag_height:n \{\textit{flag name}\}

Expands to the height of the \textit{flag} as an integer denotation.

\flag_raise:n \{\textit{flag name}\}

The \textit{flag}'s height is increased by 1 locally.
Part XIII

The \texttt{l3prg} package

Control structures

Conditional processing in \LaTeX{} has two forms of conditional flow processing based on these states. The first form is predicate functions that turn the returned state into a boolean (true) or (false). For example, the function \texttt{\cs\_if\_free:TN} checks whether the control sequence given as its argument is free and then returns the boolean (true) or (false) values to be used in testing with \texttt{\if\_predicate:w} or in functions to be described below. The second form is the kind of functions choosing a particular argument from the input stream based on the result of the testing as in \texttt{\cs\_if\_free:NTF} which also takes one argument (the N) and then executes either true or false depending on the result.

\textbf{Texhackers note:} The arguments are executed after exiting the underlying \texttt{\if...\fi:} structure.

1 Defining a set of conditional functions

\begin{verbatim}
\prg_new_conditional:Npnn \prg_new_conditional:Nnn \prg_new_protected_conditional:Npnn \prg_new_protected_conditional:Nnn
\prg_set_conditional:Npnn \prg_set_conditional:Nnn \prg_set_protected_conditional:Npnn \prg_set_protected_conditional:Nnn
\end{verbatim}

These functions create a family of conditionals using the same \{\texttt{\textlangle code\textrangle}\} to perform the test created. Those conditionals are expandable if \{\texttt{\textlangle code\textrangle}\} is. The new versions check for existing definitions and perform assignments globally (cf. \texttt{\cs\_new:Npnn}) whereas the set versions do no check and perform assignments locally (cf. \texttt{\cs\_set:Npnn}). The conditionals created are dependent on the comma-separated list of \{conditions\}, which should be one or more of p, T, F and TF.

\begin{verbatim}
\prg_new_protected_conditional:Npnn \prg_new_protected_conditional:Nnn \prg_new_protected_conditional:Nnn \prg_new_protected_conditional:Nnn
\end{verbatim}

These functions create a family of protected conditionals using the same \{\texttt{\textlangle code\textrangle}\} to perform the test created. The \{\texttt{\textlangle code\textrangle}\} does not need to be expandable. The new version check for existing definitions and perform assignments globally (cf. \texttt{\cs\_new:Npnn}) whereas the set version do not (cf. \texttt{\cs\_set:Npnn}). The conditionals created are dependent on the comma-separated list of \{conditions\}, which should be one or more of T, F and TF (not p).

The conditionals are defined by \texttt{\prg\_new\_conditional:Npnn} and friends as:
• \(\langle \text{name}\rangle_p: \langle \text{arg spec}\rangle\) — a predicate function which will supply either a logical true or logical false. This function is intended for use in cases where one or more logical tests are combined to lead to a final outcome. This function cannot be defined for protected conditionals.

• \(\langle \text{name}\rangle: \langle \text{arg spec}\rangle T\) — a function with one more argument than the original \(\langle \text{arg spec}\rangle\) demands. The \(\langle \text{true branch}\rangle\) code in this additional argument will be left on the input stream only if the test is true.

• \(\langle \text{name}\rangle: \langle \text{arg spec}\rangle F\) — a function with one more argument than the original \(\langle \text{arg spec}\rangle\) demands. The \(\langle \text{false branch}\rangle\) code in this additional argument will be left on the input stream only if the test is false.

• \(\langle \text{name}\rangle: \langle \text{arg spec}\rangle TF\) — a function with two more arguments than the original \(\langle \text{arg spec}\rangle\) demands. The \(\langle \text{true branch}\rangle\) code in the first additional argument will be left on the input stream if the test is true, while the \(\langle \text{false branch}\rangle\) code in the second argument will be left on the input stream if the test is false.

The \(\langle \text{code}\rangle\) of the test may use \(\langle \text{parameters}\rangle\) as specified by the second argument to \texttt{\prg_set_conditional:Npnn}: this should match the \(\langle \text{argument specification}\rangle\) but this is not enforced. The \texttt{\Nnn} versions infer the number of arguments from the argument specification given (cf. \texttt{\cs_new:Nn}, etc.). Within the \(\langle \text{code}\rangle\), the functions \texttt{\prg_return_true:} and \texttt{\prg_return_false:} are used to indicate the logical outcomes of the test.

An example can easily clarify matters here:

\[
\prg_set_conditional:Npnn \texttt{\foo_if_bar:NN} \{ p , T , TF \}
\{
\begin{align*}
&\text{\if_meaning:w \l_tmpa_tl #1} \\
&\text{\prg_return_true:} \\
&\text{\else:} \\
&\text{\if_meaning:w \l_tmpa_tl #2} \\
&\text{\prg_return_true:} \\
&\text{\else:} \\
&\text{\prg_return_false:} \\
&\text{\fi:} \\
&\text{\fi:}
\end{align*}
\]

This defines the function \texttt{\foo_if_bar_p:NN}, \texttt{\foo_if_bar:NNTF} and \texttt{\foo_if_bar:NNT} but not \texttt{\foo_if_bar:NNF} (because F is missing from the \(\langle \text{conditions}\rangle\) list). The return statements take care of resolving the remaining \texttt{\else:} and \texttt{\fi:} before returning the state. There must be a return statement for each branch; failing to do so will result in erroneous output if that branch is executed.

\[
\texttt{\prg_new_eq_conditional:NNn} \langle \text{name}\rangle_1: \langle \text{arg spec}\rangle_1 \langle \text{name}\rangle_2: \langle \text{arg spec}\rangle_2
\]

These functions copy a family of conditionals. The \texttt{\new} version checks for existing definitions (cf. \texttt{\cs_new_eq:NN}) whereas the \texttt{\set} version does not (cf. \texttt{\cs_set_eq:NN}). The conditionals copied are depended on the comma-separated list of \(\langle \text{conditions}\rangle\), which should be one or more of p, T, F and TF.
These “return” functions define the logical state of a conditional statement. They appear within the code for a conditional function generated by \texttt{\textbackslash prg\_set\_conditional:Npnn}, etc, to indicate when a true or false branch should be taken. While they may appear multiple times each within the code of such conditionals, the execution of the conditional must result in the expansion of one of these two functions exactly once.

The return functions trigger what is internally an f-expansion process to complete the evaluation of the conditional. Therefore, after \texttt{\textbackslash prg\_return\_true:} or \texttt{\textbackslash prg\_return\_false:} there must be no non-expandable material in the input stream for the remainder of the expansion of the conditional code. This includes other instances of either of these functions.

\texttt{\textbackslash prg\_generate\_conditional\_variant:Nnn} \hspace{1cm} \texttt{\textbackslash prg\_generate\_conditional\_variant:Nnn} \langle\text{name}\rangle:\langle\text{arg spec}\rangle \{\langle\text{variant argument specifiers}\rangle\} \{\langle\text{condition specifiers}\rangle\}

Defines argument-specifier variants of conditionals. This is equivalent to running \texttt{\textbackslash cs\_generate\_variant:Nn} \langle\text{conditional}\rangle \{\langle\text{variant argument specifiers}\rangle\} on each \langle\text{conditional}\rangle described by the \langle\text{condition specifiers}\rangle. These base-form \langle\text{conditionals}\rangle are obtained from the \langle\text{name}\rangle and \langle\text{arg spec}\rangle as described for \texttt{\textbackslash prg\_new\_conditional:Npnn}, and they should be defined.

\section{The boolean data type}

This section describes a boolean data type which is closely connected to conditional processing as sometimes you want to execute some code depending on the value of a switch \textit{e.g.}, draft/final and other times you perhaps want to use it as a predicate function in an \texttt{\textbackslash if\_predicate:w} test. The problem of the primitive \texttt{\textbackslash if\_false:} and \texttt{\textbackslash if\_true:} tokens is that it is not always safe to pass them around as they may interfere with scanning for termination of primitive conditional processing. Therefore, we employ two canonical booleans: \texttt{\textbackslash c\_true\_bool} or \texttt{\textbackslash c\_false\_bool}. Besides preventing problems as described above, it also allows us to implement a simple boolean parser supporting the logical operations And, Or, Not, \textit{etc.} which can then be used on both the boolean type and predicate functions.

All conditional \texttt{\textbackslash bool\_} functions except assignments are expandable and expect the input to also be fully expandable (which generally means being constructed from predicate functions and booleans, possibly nested).

\textbf{\textit{\texttt{\textbackslash bool\_new:N}}} The \texttt{\textbackslash bool} data type is not implemented using the \texttt{\textbackslash if\_false} / \texttt{\textbackslash if\_true} primitives, in contrast to \texttt{\textbackslash newif, etc.}, in plain \texttt{\texttt{\textbackslash LaTeX}}, \texttt{\texttt{\textbackslash LaTeX2ε}} and so on. Programmers should not base use of \texttt{\textbackslash bool} switches on any particular expectation of the implementation.

\texttt{\textbackslash bool\_new:N} \langle\text{boolean}\rangle

Creates a new \langle\text{boolean}\rangle or raises an error if the name is already taken. The declaration is global. The \langle\text{boolean}\rangle is initially \texttt{\textbackslash texttt{false}}.
\bool_const:Nn
\bool_const:cn

\bool_set_false:N (boolean)
\bool_set_false:c
\bool_gset_false:N
\bool_gset_false:c

\bool_set_true:N (boolean)
\bool_set_true:c
\bool_gset_true:N
\bool_gset_true:c

\bool_set_eq:NN ⟨boolean 1⟩ ⟨boolean 2⟩
\bool_set_eq:cn (cN|Nc|cc)
\bool_gset_eq:NN
\bool_gset_eq:cc (cN|Nc|cc)

\bool_set:Nn
\bool_set:cn
\bool_gset:Nn
\bool_gset:cn

\bool_if_p:N ⋆
\bool_if_p:c ⋆
\bool_if:N ⋆
\bool_if:c ⋆

\bool_show:N
\bool_show:c

\bool_show:n

\bool_log:N
\bool_log:c

\bool_const:Nn (boolean) \langle\text{boolexpr}\rangle

Creates a new constant \langle boolean \rangle or raises an error if the name is already taken. The value of the \langle boolean \rangle is set globally to the result of evaluating the \langle boolexpr \rangle.

\bool_set_false:N (boolean)

Sets \langle boolean \rangle logically false.

\bool_set_true:N (boolean)

Sets \langle boolean \rangle logically true.

\bool_set_eq:NN \langle boolean 1⟩ \langle boolean 2⟩

Sets \langle boolean 1⟩ to the current value of \langle boolean 2⟩.

\bool_set:Nn (boolean) \langle\text{boolexpr}\rangle

Evaluates the \langle boolean expression \rangle as described for \bool_if:nTF, and sets the \langle boolean \rangle variable to the logical truth of this evaluation.

\bool_if_p:N ⋆
\bool_if_p:c ⋆
\bool_if:N ⋆
\bool_if:c ⋆

Tests the current truth of \langle boolean \rangle, and continues expansion based on this result.

\bool_show:N (boolean)

Displays the logical truth of the \langle boolean \rangle on the terminal.

\bool_show:n \langle boolean expression⟩

Displays the logical truth of the \langle boolean expression \rangle on the terminal.

\bool_log:N (boolean)

Writes the logical truth of the \langle boolean \rangle in the log file.
\bool_log:n { (boolean expression) }
Wrightes the logical truth of the \langle boolean expression \rangle in the log file.

\bool_if_exist_p:N (boolean)
\bool_if_exist:NTF (boolean) \{ (true code) \} \{ (false code) \}
Tests whether the \langle boolean \rangle is currently defined. This does not check that the \langle boolean \rangle really is a boolean variable.

\bool_if_exist_p:c ⋆
\bool_if_exist_p:N ⋆
\bool_if_exist:N TF ⋆
\bool_if_exist:c TF ⋆
New: 2012-03-03

\l_tmpa_bool \l_tmpb_bool

\g_tmpa_bool \g_tmpb_bool

A scratch boolean for local assignment. It is never used by the kernel code, and so is safe for use with any \LaTeX3-defined function. However, it may be overwritten by other non-kernel code and so should only be used for short-term storage.

A scratch boolean for global assignment. It is never used by the kernel code, and so is safe for use with any \LaTeX3-defined function. However, it may be overwritten by other non-kernel code and so should only be used for short-term storage.

3 Boolean expressions

As we have a boolean datatype and predicate functions returning boolean \langle true \rangle or \langle false \rangle values, it seems only fitting that we also provide a parser for \langle boolean expressions \rangle.

A boolean expression is an expression which given input in the form of predicate functions and boolean variables, return boolean \langle true \rangle or \langle false \rangle. It supports the logical operations And, Or and Not as the well-known infix operators \&\& and || and prefix ! with their usual precedences (namely, \&\& binds more tightly than ||). In addition to this, parentheses can be used to isolate sub-expressions. For example,

\int_compare_p:n { 1 = 1 } \&\&
( \int_compare_p:n { 2 = 3 } ||
\int_compare_p:n { 4 \leq 4 } ||
\str_if_eq_p:nn { abc } { def } ) \&\&
! \int_compare_p:n { 2 = 4 }

is a valid boolean expression.

Contrarily to some other programming languages, the operators \&\& and || evaluate both operands in all cases, even when the first operand is enough to determine the result. This “eager” evaluation should be contrasted with the “lazy” evaluation of \bool_lazy_- functions.

\TeXhackers note: The eager evaluation of boolean expressions is unfortunately necessary in \TeX. Indeed, a lazy parser can get confused if \&\& or || or parentheses appear as (unbraced) arguments of some predicates. For instance, the innocuous-looking expression below would break (in a lazy parser) if \#1 were a closing parenthesis and \l_tmpa_bool were true.

( \l_tmpa_bool || \token_if_eq_meaning_p:NN X \#1 )
Minimal (lazy) evaluation can be obtained using the conditionals \bool_lazy_all:nTF, \bool_lazy_and:nnTF, \bool_lazy_any:nTF, or \bool_lazy_or:nnTF, which only evaluate their boolean expression arguments when they are needed to determine the resulting truth value. For example, when evaluating the boolean expression

\bool_lazy_and_p:nn
{
  \bool_lazy_any_p:n
  {
  \int_compare_p:n { 2 = 3 }
  \int_compare_p:n { 4 \leq 4 }
  \int_compare_p:n { 1 = \error } } % skipped
  }
  
  { ! \int_compare_p:n { 2 = 4 } }

the line marked with \texttt{skipped} is not expanded because the result of \bool_lazy_any_p:n is known once the second boolean expression is found to be logically true. On the other hand, the last line is expanded because its logical value is needed to determine the result of \bool_lazy_and_p:nn.

\begin{verbatim}
\bool_if_p:n\bool_if:nTF

Tests the current truth of \langle boolean expression \rangle, and continues expansion based on this result. The \langle boolean expression \rangle should consist of a series of predicates or boolean variables with the logical relationship between these defined using && ("And"), || ("Or"), ! ("Not") and parentheses. The logical Not applies to the next predicate or group.

\begin{verbatim}
\bool_lazy_all_p:n\bool_lazy_all:nTF

Implements the "And" operation on the \langle boolean expressions \rangle, hence is true if all of them are true and false if any of them is false. Contrarily to the infix operator &&, only the \langle boolean expressions \rangle which are needed to determine the result of \bool_lazy_all:nTF are evaluated. See also \bool_lazy_and:nnTF when there are only two \langle boolean expressions \rangle.

\begin{verbatim}
\bool_lazy_and_p:nn\bool_lazy_and:nnTF

Implements the "And" operation between two boolean expressions, hence is true if both are true. Contrarily to the infix operator \&, the \langle boolexpr2 \rangle is only evaluated if it is needed to determine the result of \bool_lazy_and:nnTF. See also \bool_lazy_all:nTF when there are more than two \langle boolean expressions \rangle.
\begin{itemize}
\item \texttt{\bool_lazy_or_p:n} \* \item \texttt{\bool_lazy_or_p nn} \* \\
\texttt{\bool_lazy_or_p:nTF} \* \item \texttt{\bool_lazy_or_p nnTF} \* \\
\texttt{\bool_lazy_or_p:nTF} \* \item \texttt{\bool_lazy_or_p:nTF} \* \\
\item \texttt{\bool_lazy_any:p:n} \* \item \texttt{\bool_lazy_any:nTF} \* \\
\item \texttt{\bool_lazy_any:p:nTF} \* \item \texttt{\bool_lazy_any:nTF} \* \\
\item \texttt{\bool_lazy_or:p:n} \* \item \texttt{\bool_lazy_or:p nn} \* \\
\item \texttt{\bool_lazy_or:p:nTF} \* \item \texttt{\bool_lazy_or:p nnTF} \* \\
\item \texttt{\bool_lazy_or:p:nTF} \* \item \texttt{\bool_lazy_or:p:nTF} \* \\

\end{itemize}

\begin{itemize}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\item \texttt{\bool_lazy_or:p:n} \{ \langle boolean expression\rangle \} \{ \langle boolean expression\rangle \} \langle boolean expression\rangle \ldots \{ \langle boolean expression\rangle \} \{ \langle true code\rangle \} \{ \langle false code\rangle \}
\end{itemize}

\section*{4 Logical loops}

Loops using either boolean expressions or stored boolean values.

\begin{itemize}
\item \texttt{\bool_do_until:Nn} \* \item \texttt{\bool_do_until:cn} \* \\
\texttt{\bool_do_until:Nn} \* \item \texttt{\bool_do_until:cn} \* \\
\item \texttt{\bool_until_do:Nn} \* \item \texttt{\bool_until_do:cn} \* \\
\item \texttt{\bool_until_do:Nn} \* \item \texttt{\bool_until_do:cn} \* \\
\item \texttt{\bool_while_do:Nn} \* \item \texttt{\bool_while_do:cn} \* \\
\item \texttt{\bool_while_do:Nn} \* \item \texttt{\bool_while_do:cn} \* \\
\end{itemize}

\begin{itemize}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_do_until:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\end{itemize}

\begin{itemize}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\end{itemize}

\begin{itemize}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\end{itemize}

\begin{itemize}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_until_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\end{itemize}

\begin{itemize}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\item \texttt{\bool_while_do:Nn} \{ \langle boolean expression\rangle \} \{ \langle code\rangle \}
\end{itemize}
\bool_do_until:nn \{\text{boolean expression}\} \{\text{code}\}

Places the \{\text{code}\} in the input stream for \TeX{} to process, and then checks the logical value of the \{\text{boolean expression}\} as described for \bool_if:nTF. If it is \texttt{false} then the \{\text{code}\} is inserted into the input stream again and the process loops until the \{\text{boolean expression}\} evaluates to \texttt{true}.

\bool_do_while:nn \{\text{boolean expression}\} \{\text{code}\}

Places the \{\text{code}\} in the input stream for \TeX{} to process, and then checks the logical value of the \{\text{boolean expression}\} as described for \bool_if:nTF. If it is \texttt{true} then the \{\text{code}\} is inserted into the input stream again and the process loops until the \{\text{boolean expression}\} evaluates to \texttt{false}.

\bool_until_do:nn \{\text{boolean expression}\} \{\text{code}\}

This function firsts checks the logical value of the \{\text{boolean expression}\} (as described for \bool_if:nTF). If it is \texttt{false} the \{\text{code}\} is placed in the input stream and expanded. After the completion of the \{\text{code}\} the truth of the \{\text{boolean expression}\} is re-evaluated. The process then loops until the \{\text{boolean expression}\} is \texttt{true}.

\bool_while_do:nn \{\text{boolean expression}\} \{\text{code}\}

This function firsts checks the logical value of the \{\text{boolean expression}\} (as described for \bool_if:nTF). If it is \texttt{true} the \{\text{code}\} is placed in the input stream and expanded. After the completion of the \{\text{code}\} the truth of the \{\text{boolean expression}\} is re-evaluated. The process then loops until the \{\text{boolean expression}\} is \texttt{false}.

\prg_replicate:nn \{\text{integer expression}\} \{\text{tokens}\}

Evaluates the \{\text{integer expression}\} (which should be zero or positive) and creates the resulting number of copies of the \{\text{tokens}\}. The function is both expandable and safe for nesting. It yields its result after two expansion steps.

5 Producing multiple copies

6 Detecting \TeX{}’s mode

\mode_if_horizontal_p: \* \mode_if_horizontal:TF \{\texttt{true code}\} \{\texttt{false code}\}

Detects if \TeX{} is currently in horizontal mode.

\mode_if_inner_p: \* \mode_if_inner:TF \{\texttt{true code}\} \{\texttt{false code}\}

Detects if \TeX{} is currently in inner mode.

\mode_if_math_p: \* \mode_if_math:TF \{\texttt{true code}\} \{\texttt{false code}\}

Detects if \TeX{} is currently in maths mode.

Updated: 2011-07-15

Updated: 2017-07-15

Updated: 2011-07-04

Updated: 2011-09-05
Detects if \TeX{} is currently in vertical mode.

7 Primitive conditionals

\texttt{\if_predicate:w \langle predicate \rangle \langle true \ code \rangle \else: \langle false \ code \rangle \fi:}

This function takes a predicate function and branches according to the result. (In practice this function would also accept a single boolean variable in place of the \langle predicate \rangle but to make the coding clearer this should be done through \texttt{\if_bool:N}.)

\texttt{\if_bool:N \langle boolean \rangle \langle true \ code \rangle \else: \langle false \ code \rangle \fi:}

This function takes a boolean variable and branches according to the result.

8 Nestable recursions and mappings

There are a number of places where recursion or mapping constructs are used in \texttt{expl3}. At a low-level, these typically require insertion of tokens at the end of the content to allow “clean up”. To support such mappings in a nestable form, the following functions are provided.

\texttt{\prg_break_point:Nn \langle type \rangle _map_break: \{ \langle code \rangle \}}

Used to mark the end of a recursion or mapping: the functions \texttt{\langle type \rangle _map_break:} and \texttt{\langle type \rangle _map_break:n} use this to break out of the loop (see \texttt{\prg_map_break:Nn} for how to set these up). After the loop ends, the \langle code \rangle is inserted into the input stream. This occurs even if the break functions are not applied: \texttt{\prg_break_point:Nn} is functionally-equivalent in these cases to \texttt{\use_ii:nn}.

\texttt{\prg_map_break:Nn \langle type \rangle _map_break: \{ \langle user \ code \rangle \}}

\texttt{\prg_break_point:Nn \langle type \rangle _map_break: \{ \langle ending \ code \rangle \}}

Breaks a recursion in mapping contexts, inserting in the input stream the \langle user \ code \rangle after the \langle ending \ code \rangle for the loop. The function breaks loops, inserting their \langle ending \ code \rangle, until reaching a loop with the same \langle type \rangle as its first argument. This \texttt{\langle type \rangle _map_break:} argument must be defined; it is simply used as a recognizable marker for the \langle type \rangle.

For types with mappings defined in the kernel, \texttt{\langle type \rangle _map_break:} and \texttt{\langle type \rangle _map_break:n} are defined as \texttt{\prg_map_break:Nn \langle type \rangle _map_break: \{ \} } and the same with \{ \} omitted.

8.1 Simple mappings

In addition to the more complex mappings above, non-nestable mappings are used in a number of locations and support is provided for these.
This copy of `\prg_do_nothing:` is used to mark the end of a fast short-term recursion: the function `\prg_break:n` uses this to break out of the loop.

```
\prg_break:n \langle code \rangle \ldots \prg_break_point:
```

Breaks a recursion which has no \emph{ending code} and which is not a user-breakable mapping (see for instance `\prop_get:Nn`), and inserts the \emph{code} in the input stream.

9 Internal programming functions

```
\group_align_safe_begin: ...
\group_align_safe_end: ...
```

These functions are used to enclose material in a \TeX{} alignment environment within a specially-constructed group. This group is designed in such a way that it does not add brace groups to the output but does act as a group for the \& token inside `\halign`. This is necessary to allow grabbing of tokens for testing purposes, as \TeX{} uses group level to determine the effect of alignment tokens. Without the special grouping, the use of a function such as `\peek_after:Nw` would result in a forbidden comparison of the internal `\endtemplate` token, yielding a fatal error. Each `\group_align_safe_begin:` must be matched by a `\group_align_safe_end;`, although this does not have to occur within the same function.
Part XIV
The l3sys package: System/runtime functions

1 The name of the job

\c_sys_jobname_str

Constant that gets the “job name” assigned when \TeX starts.

\TeXhackers note: This copies the contents of the primitive \jobname. For technical reasons, the string here is not of the same internal form as other, but may be manipulated using normal string functions.

2 Date and time

\c_sys_minute_int
\c_sys_hour_int
\c_sys_day_int
\c_sys_month_int
\c_sys_year_int

The date and time at which the current job was started: these are all reported as integers.

\TeXhackers note: Whilst the underlying primitives can be altered by the user, this interface to the time and date is intended to be the “real” values.

3 Engine

\sys_if_engine_luatex_p: *
\sys_if_engine_pdftex_p: *
\sys_if_engine_ptex_p: *
\sys_if_engine_uptex_p: *
\sys_if_engine_xetex_p: *
\sys_if_engine_luatex:TF {⟨true code⟩} {⟨false code⟩}

Conditional which allow engine-specific code to be used. The names follow naturally from those of the engine binaries: note that the (u)ptex tests are for ε-\TeX and ε-up\TeX as expl3 requires the ε-\TeX extensions. Each conditional is true for exactly one supported engine. In particular, \sys_if_engine_ptex_p: is true for ε-\TeX but false for ε-up\TeX.

\c_sys_engine_str

The current engine given as a lower case string: one of luatex, pdftex, ptex, uptex or xetex.
4 Output format

\sys_if_output_dvi_p: \* \sys_if_output_dvi:TF \{\langle true code\rangle\} \{\langle false code\rangle\}

Conditionals which give the current output mode the \TeX run is operating in. This is always one of two outcomes, DVI mode or PDF mode. The two sets of conditionals are thus complementary and are both provided to allow the programmer to emphasise the most appropriate case.

\c_sys_output_str

New: 2015-09-19

The current output mode given as a lower case string: one of dvi or pdf.

5 Platform

\sys_if_platform_unix_p: \* \sys_if_platform_unix:TF \{\langle true code\rangle\} \{\langle false code\rangle\}

\sys_if_platform_windows_p: \* \sys_if_platform_windows:TF \* \sys_if_platform_unix:TF \* \sys_if_platform_windows:TF \*

Conditionals which allow platform-specific code to be used. The names follow the Lua os.type() function, i.e. all Unix-like systems are unix (including Linux and MacOS).

\c_sys_platform_str

New: 2018-07-27

The current platform given as a lower case string: one of unix, windows or unknown.

6 Random numbers

\sys_rand_seed:

Expands to the current value of the engine’s random seed, a non-negative integer. In engines without random number support this expands to 0.

\sys_gset_rand_seed:n \{\langle intexpr\rangle\}

New: 2017-05-27

Globally sets the seed for the engine’s pseudo-random number generator to the \langle integer expression\rangle. This random seed affects all \{\ldots\_rand functions (such as \int_rand:nn or \clist_rand_item:n) as well as other packages relying on the engine’s random number generator. In engines without random number support this produces an error.

\TeXhackers note: While a 32-bit (signed) integer can be given as a seed, only the absolute value is used and any number beyond $2^{28}$ is divided by an appropriate power of 2. We recommend using an integer in $[0, 2^{28} - 1]$. 
7 Access to the shell

\sys_get_shell:nnN \sys_get_shell:nnN \{\langle shell command\rangle\} \{\langle setup\rangle\} \{tl var\} \sys_get_shell:nnN \{\langle shell command\rangle\} \{\langle setup\rangle\} \{tl var\} \{\{true code\}\} \{\{false code\}\}

Defines \{tl\} to the text returned by the \{\langle shell command\rangle\}. The \{\langle shell command\rangle\} is converted to a string using \tl_to_str:n. Category codes may need to be set appropriately via the \{\langle setup\rangle\} argument, which is run just before running the \{\langle shell command\rangle\} (in a group). If shell escape is disabled, the \{tl var\} will be set to \q_no_value in the non-branching version. Note that quote characters (" cannot be used inside the \{\langle shell command\rangle\}. The \sys_get_shell:nnNTF conditional returns true if the shell is available and no quote is detected, and false otherwise.

\c_sys_shell_escape_int

This variable exposes the internal triple of the shell escape status. The possible values are:

0 Shell escape is disabled
1 Unrestricted shell escape is enabled
2 Restricted shell escape is enabled

\sys_if_shell:p: \sys_if_shell:TF:
\sys_if_shell_unrestricted:p: \sys_if_shell_unrestricted:TF:
\sys_if_shell_restricted:p: \sys_if_shell_restricted:TF:

Performs a check for whether shell escape is enabled. This returns true if either of restricted or unrestricted shell escape is enabled.

Performs a check for whether unrestricted shell escape is enabled.

Performs a check for whether restricted shell escape is enabled. This returns false if unrestricted shell escape is enabled. Unrestricted shell escape is not considered a superset of restricted shell escape in this case. To find whether any shell escape is enabled use \sys_if_shell:..\nsys_shell_now:n \sys_shell_now:x
\sys_shell_shipout:n \sys_shell_shipout:x

Execute \{tokens\} through shell escape immediately.

Execute \{tokens\} through shell escape at shipout.
7.1 Loading configuration data

\sys_load_backend:n {\textit{backend}}

Loads the additional configuration file needed for backend support. If the \textit{backend} is empty, the standard backend for the engine in use will be loaded. This command may only be used once.

\c_sys_backend_str

Set to the name of the backend in use by \sys_load_backend:n when issued.

\sys_load_debug:
\sys_load_deprecation:

Load the additional configuration files for debugging support and rolling back deprecations, respectively.

7.2 Final settings

\sys_finalise:

Finalises all system-dependent functionality: required before loading a backend.
Part XV

The \texttt{l3clist} package

Comma separated lists

Comma lists contain ordered data where items can be added to the left or right end of the list. This data type allows basic list manipulations such as adding/removing items, applying a function to every item, removing duplicate items, extracting a given item, using the comma list with specified separators, and so on. Sequences (defined in \texttt{l3seq}) are safer, faster, and provide more features, so they should often be preferred to comma lists. Comma lists are mostly useful when interfacing with \LaTeX\ or other code that expects or provides comma list data.

Several items can be added at once. To ease input of comma lists from data provided by a user outside an \texttt{\ExplSyntaxOn ... \ExplSyntaxOff} block, spaces are removed from both sides of each comma-delimited argument upon input. Blank arguments are ignored, to allow for trailing commas or repeated commas (which may otherwise arise when concatenating comma lists “by hand”). In addition, a set of braces is removed if the result of space-trimming is braced: this allows the storage of any item in a comma list. For instance,

\begin{verbatim}
\clist_new:N \l_my_clist
\clist_put_left:Nn \l_my_clist { -a- , -{b}-, c\d }
\clist_put_right:Nn \l_my_clist { -{e-} , , {f} , }
\end{verbatim}

results in \texttt{\l_my_clist} containing \texttt{a,b,c-d, e-} and \texttt{f}. Comma lists normally do not contain empty items so the following gives an empty comma list:

\begin{verbatim}
\clist_clear_new:N \l_my_clist
\clist_put_right:Nn \l_my_clist { , , , }
\clist_if_empty:NTF \l_my_clist { true } { false }
\end{verbatim}

and it leaves \texttt{true} in the input stream. To include an “unsafe” item (empty, or one that contains a comma, or starts or ends with a space, or is a single brace group), surround it with braces.

Almost all operations on comma lists are noticeably slower than those on sequences so converting the data to sequences using \texttt{\seq_set_from_clist:Nn} (see \texttt{l3seq}) may be advisable if speed is important. The exception is that \texttt{\clist_if_in:NnTF} and \texttt{\clist_remove_duplicates:N} may be faster than their sequence analogues for large lists. However, these functions work slowly for “unsafe” items that must be braced, and may produce errors when their argument contains \{, \} or \# (assuming the usual \TeX\ category codes apply). In addition, comma lists cannot store quarks \texttt{\q_mark} or \texttt{\q_stop}. The sequence data type should thus certainly be preferred to comma lists to store such items.

\section{Creating and initialising comma lists}

\texttt{\clist_new:N \clist_new:c}

Creates a new \texttt{(comma list)} or raises an error if the name is already taken. The declaration is global. The \texttt{(comma list)} initially contains no items.
\clist_const:Nn \clist_const:(Nx|cn|cx) \new: 2014-07-05

Creates a new constant ⟨clist var ⟩ or raises an error if the name is already taken. The value of the ⟨clist var ⟩ is set globally to the ⟨comma list ⟩.

\clist_clear:N \clist_clear:c \clist_gclear:N \clist_gclear:c

Clears all items from the ⟨comma list ⟩.

\clist_clear_new:N \clist_clear_new:c \clist_gclear_new:N \clist_gclear_new:c \new: 2014-07-17

Ensures that the ⟨comma list ⟩ exists globally by applying \clist_new:N if necessary, then applies \clist_(g)clear:N to leave the list empty.

\clist_set_eq:NN \clist_set_eq:(cN|Nc|cc) \clist_gset_eq:NN \clist_gset_eq:(cN|Nc|cc) \new: 2012-03-03

Sets the content of ⟨comma list1 ⟩ equal to that of ⟨comma list2 ⟩.

\clist_set_from_seq:NN \clist_set_from_seq:(cN|Nc|cc) \clist_gset_from_seq:NN \clist_gset_from_seq:(cN|Nc|cc) \new: 2014-07-17

Converts the data in the ⟨sequence ⟩ into a ⟨comma list ⟩: the original ⟨sequence ⟩ is unchanged. Items which contain either spaces or commas are surrounded by braces.

\clist_concat:NNN \clist_concat:ccc \clist_gconcat:NNN \clist_gconcat:ccc

Concatenates the content of ⟨comma list1 ⟩ and ⟨comma list2 ⟩ together and saves the result in ⟨comma list1 ⟩. The items in ⟨comma list2 ⟩ are placed at the left side of the new comma list.

\clist_if_exist_p:N \clist_if_exist_p:c \clist_if_exist:NTF \clist_if_exist:CTF \new: 2012-03-03

Tests whether the ⟨comma list ⟩ is currently defined. This does not check that the ⟨comma list ⟩ really is a comma list.
2 Adding data to comma lists

\clist_set:Nn ⟨ comma list ⟩ \{ ⟨ item ⟩ ,...,⟨ item ⟩ \}
\clist_set:Nn \clist_set: ⟨ NV|No|Nx|cn|cV|co|cx ⟩
\clist_gset:Nn \clist_gset: ⟨ NV|No|Nx|cn|cV|co|cx ⟩

Sets ⟨ comma list ⟩ to contain the ⟨ items ⟩, removing any previous content from the variable. Blank items are omitted, spaces are removed from both sides of each item, then a set of braces is removed if the resulting space-trimmed item is braced. To store some ⟨ tokens ⟩ as a single ⟨ item ⟩ even if the ⟨ tokens ⟩ contain commas or spaces, add a set of braces:
\clist_set:Nn \clist_set: ⟨ comma list ⟩ \{ { ⟨ tokens ⟩ } \}.

\clist_put_left:Nn \clist_put_left: ⟨ comma list ⟩ \{ ⟨ item ⟩ ,...,⟨ item ⟩ \}
\clist_put_left:Nn \clist_put_left: ⟨ NV|No|Nx|cn|cV|co|cx ⟩

Appends the ⟨ items ⟩ to the left of the ⟨ comma list ⟩. Blank items are omitted, spaces are removed from both sides of each item, then a set of braces is removed if the resulting space-trimmed item is braced. To append some ⟨ tokens ⟩ as a single ⟨ item ⟩ even if the ⟨ tokens ⟩ contain commas or spaces, add a set of braces:
\clist_put_left:Nn \clist_put_left: ⟨ comma list ⟩ \{ { ⟨ tokens ⟩ } \}.

\clist_put_right:Nn \clist_put_right: ⟨ comma list ⟩ \{ ⟨ item ⟩ ,...,⟨ item ⟩ \}
\clist_put_right:Nn \clist_put_right: ⟨ NV|No|Nx|cn|cV|co|cx ⟩

Appends the ⟨ items ⟩ to the right of the ⟨ comma list ⟩. Blank items are omitted, spaces are removed from both sides of each item, then a set of braces is removed if the resulting space-trimmed item is braced. To append some ⟨ tokens ⟩ as a single ⟨ item ⟩ even if the ⟨ tokens ⟩ contain commas or spaces, add a set of braces:
\clist_put_right:Nn \clist_put_right: ⟨ comma list ⟩ \{ { ⟨ tokens ⟩ } \}.

3 Modifying comma lists

While comma lists are normally used as ordered lists, it may be necessary to modify the content. The functions here may be used to update comma lists, while retaining the order of the unaffected entries.
\clist_remove_duplicates:N \clist_remove_duplicates:N ⟨comma list⟩
\clist_remove_duplicates:c
\clist_gremove_duplicates:N
\clist_gremove_duplicates:c

Removes duplicate items from the ⟨comma list⟩, leaving the left most copy of each item in the ⟨comma list⟩. The ⟨item⟩ comparison takes place on a token basis, as for \tl_if_eq:nn(TF).

\textbf{\TeXhackers note:} This function iterates through every item in the ⟨comma list⟩ and does a comparison with the ⟨items⟩ already checked. It is therefore relatively slow with large comma lists. Furthermore, it may fail if any of the items in the ⟨comma list⟩ contains ⟨,⟩, or # (assuming the usual \TeX category codes apply).

\clist_remove_all:Nn \clist_remove_all:N ⟨comma list⟩ {⟨item⟩}
\clist_remove_all:cn
\clist_gremove_all:Nn \clist_gremove_all:cn

Removes every occurrence of ⟨item⟩ from the ⟨comma list⟩. The ⟨item⟩ comparison takes place on a token basis, as for \tl_if_eq:nn(TF).

\textbf{\TeXhackers note:} The function may fail if the ⟨item⟩ contains ⟨,⟩, or # (assuming the usual \TeX category codes apply).

\clist_reverse:N \clist_reverse:N ⟨comma list⟩
\clist_reverse:c
\clist_greverse:N \clist_greverse:c

Reverses the order of items stored in the ⟨comma list⟩.

\textbf{\TeXhackers note:} The result is returned within \unexpanded, which means that the comma list does not expand further when appearing in an x-type or e-type argument expansion.

\clist_reverse:n \clist_reverse:n ⟨comma list⟩

Leaves the items in the ⟨comma list⟩ in the input stream in reverse order. Contrarily to other what is done for other \n-type ⟨comma list⟩ arguments, braces and spaces are preserved by this process.

\textbf{\TeXhackers note:} The result is returned within \unexpanded, which means that the comma list does not expand further when appearing in an x-type or e-type argument expansion.

\clist_sort:Nn \clist_sort:Nn ⟨clist var⟩ {⟨comparison code⟩}
\clist_sort:cn
\clist_gsort:Nn \clist_gsort:cn
\textbf{\TeXhackers note:} This function sorts the ⟨clist var⟩ according to the ⟨comparison code⟩, and assigns the result to ⟨clist var⟩. The details of sorting comparison are described in Section 1.

Updated: 2011–09–06
New: 2014–07–18
New: 2017–02–06

121
4 Comma list conditionals

\clist_if_empty_p:N \clist_if_empty_p:NTF \clist_if_empty:NTF
\clist_if_empty_p:n \clist_if_empty:nTF
\clist_if_empty_p:c \clist_if_empty:cTF
\clist_if_empty_p:cn \clist_if_empty:cnTF

Tests if the \texttt{⟨comma list⟩} is empty (containing no items). The rules for space trimming are as for other n-type comma-list functions, hence the comma list \{~,~,~,\} (without outer braces) is empty, while \{~,\}, \} (without outer braces) contains one element, which happens to be empty: the comma-list is not empty.

\clist_if_in:NnTF \clist_if_in:TF \clist_if_in:nnTF
\clist_if_in:NTF \clist_if_in:TNF \clist_if_in:nn\texttt{T}F
\clist_if_in:nTF \clist_if_in:nnTF
\clist_if_in:cTF \clist_if_in:cnTF
\clist_if_in:cn\texttt{T}F

Tests if the \texttt{⟨item⟩} is present in the \texttt{⟨comma list⟩}. In the case of an n-type \texttt{⟨comma list⟩}, the usual rules of space trimming and brace stripping apply. Hence, \[
\clist_if_in:nnTF \{ a , \{b\}~ , \{b\} , c \} \{ b \} \{true \} \{false \}
\]
yields \texttt{true}.

\textit{T\TeX}hackers note: The function may fail if the \texttt{⟨item⟩} contains \{, \}, or \# (assuming the usual \TeX{} category codes apply).

5 Mapping to comma lists

The functions described in this section apply a specified function to each item of a comma list. All mappings are done at the current group level, \textit{i.e.} any local assignments made by the \texttt{⟨function⟩} or \texttt{⟨code⟩} discussed below remain in effect after the loop.

When the comma list is given explicitly, as an n-type argument, spaces are trimmed around each item. If the result of trimming spaces is empty, the item is ignored. Otherwise, if the item is surrounded by braces, one set is removed, and the result is passed to the mapped function. Thus, if the comma list that is being mapped is \{a,\{b\},\}, then the arguments passed to the mapped function are ‘a’, ‘\{b\}’, an empty argument, and ‘\texttt{\texttt{c}}’.

When the comma list is given as an N-type argument, spaces have already been trimmed on input, and items are simply stripped of one set of braces if any. This case is more efficient than using n-type comma lists.

\clist_map_function:NN \clist_map_function:cN \clist_map_function:nN
\clist_map_function:NN\texttt{\texttt{T}F}
\clist_map_function:cN\texttt{\texttt{T}F}
\clist_map_function:nN\texttt{\texttt{T}F}

Applies \texttt{⟨function⟩} to every \texttt{⟨item⟩} stored in the \texttt{⟨comma list⟩}. The \texttt{⟨function⟩} receives one argument for each iteration. The \texttt{⟨items⟩} are returned from left to right. The function \\texttt{clist_map_inline:Nn} is in general more efficient than \texttt{clist_map_function:NN}.
\clist_map_inline:Nn \clist_map_inline:cn \clist_map_inline:nn

Applies \texttt{(inline function)} to every \texttt{(item)} stored within the \texttt{(comma list)}. The \texttt{(inline function)} should consist of code which receives the \texttt{(item)} as \texttt{#1}. The \texttt{(items)} are returned from left to right.

\clist_map_variable:NNn \clist_map_variable:cn \clist_map_variable:nn

Stores each \texttt{(item)} of the \texttt{(comma list)} in turn in the (token list) \texttt{(variable)} and applies the \texttt{(code)}. The \texttt{(code)} will usually make use of the \texttt{(variable)}, but this is not enforced. The assignments to the \texttt{(variable)} are local. Its value after the loop is the last \texttt{(item)} in the \texttt{(comma list)}, or its original value if there were no \texttt{(item)}. The \texttt{(items)} are returned from left to right.

\clist_map_break:

Used to terminate a \texttt{\clist_map...} function before all entries in the \texttt{(comma list)} have been processed. This normally takes place within a conditional statement, for example

\begin{verbatim}
\clist_map_inline:Nn \l_my_clist
{ \str_if_eq:nnTF { #1 } { bingo } { \clist_map_break: }
  { % Do something useful
  }
}
\end{verbatim}

Use outside of a \texttt{\clist_map...} scenario leads to low level \TeX errors.

\textbf{\TeX hackers note:} When the mapping is broken, additional tokens may be inserted before further items are taken from the input stream. This depends on the design of the mapping function.
\clist_map_break:n \clist_map_break:n \{\langle code\rangle\}\n
Used to terminate a \clist_map... function before all entries in the \langle comma list \rangle have been processed, inserting the \langle code \rangle after the mapping has ended. This normally takes place within a conditional statement, for example

\clist_map_inline:Nn \l_my_clist
{\str_if_eq:nnTF { \l_\#1 } { bingo } \{ \clist_map_break:n \{ <\langle code\rangle\} \} \% Do something useful \}}\n
Use outside of a \clist_map... scenario leads to low level \TeX errors.

\TeXhackers note: When the mapping is broken, additional tokens may be inserted before the \langle code \rangle is inserted into the input stream. This depends on the design of the mapping function.

\clist_count:N \clist_count:c \clist_count:n \n
Leaves the number of items in the \langle comma list \rangle in the input stream as an \langle integer denotation \rangle. The total number of items in a \langle comma list \rangle includes those which are duplicates, i.e. every item in a \langle comma list \rangle is counted.

6 Using the content of comma lists directly

\clist_use:Nnnn \clist_use:cnnn \clist_use:nnnn \n
Places the contents of the \langle clist var \rangle in the input stream, with the appropriate \langle separator \rangle between the items. Namely, if the comma list has more than two items, the \langle separator between more than two \rangle is placed between each pair of items except the last, for which the \langle separator between final two \rangle is used. If the comma list has exactly two items, then they are placed in the input stream separated by the \langle separator between two \rangle. If the comma list has a single item, it is placed in the input stream, and a comma list with no items produces no output. An error is raised if the variable does not exist or if it is invalid.

For example,

\clist_set:Nn \l_ttmpa_clist \{ a , b , , c , {de} , f \} \clist_use:Nnnn \l_ttmpa_clist \{ -\text{-and-} \} \{ ,\text{-} \} \{ ,\text{-and-} \}

inserts "a, b, c, de, and f" in the input stream. The first separator argument is not used in this case because the comma list has more than 2 items.

\TeXhackers note: The result is returned within the \langle unexpanded \rangle primitive (\\exp_not:n), which means that the \langle items \rangle do not expand further when appearing in an x-type argument expansion.
\clist_use:Nn \clist_use:cn
\begin{itemize}
\item \clist_use:Nn \clist_var \sep\end{itemize}
Places the contents of the \clist_var in the input stream, with the \sep between the items. If the comma list has a single item, it is placed in the input stream, and a comma list with no items produces no output. An error is raised if the variable does not exist or if it is invalid.

For example,
\begin{verbatim}
\clist_set:Nn \l_tmpa_clist { a , b , , c , {de} , f }
\clist_use:Nn \l_tmpa_clist { ~and~ }
\end{verbatim}
inserts “a and b and c and de and f” in the input stream.

\textbf{\TeX hackers note:} The result is returned within the \texttt{\unexpanded} primitive (\verb|\exp_not:n|), which means that the \items do not expand further when appearing in an x-type argument expansion.

\section{Comma lists as stacks}

Comma lists can be used as stacks, where data is pushed to and popped from the top of the comma list. (The left of a comma list is the top, for performance reasons.) The stack functions for comma lists are not intended to be mixed with the general ordered data functions detailed in the previous section: a comma list should either be used as an ordered data type or as a stack, but not in both ways.

\begin{itemize}
\item \clist_get:NN \clist_pop:NN
\item \clist_gpop:NN \clist_pop:NNTF
\end{itemize}

\textbf{\TeX hackers note:} If the \clist is empty, leaves the \false code in the input stream. The value of the \token_list variable is not defined in this case and should not be relied upon. If the \clist is non-empty, pops the top item from the \clist in the \token_list variable, \i.e. removes the item from the \clist. Both the \clist and the \token_list are assigned locally.
\clist_gpop:NNTF \clist_gpop:cN

If the \langle comma list \rangle is empty, leaves the \langle false code \rangle in the input stream. The value of the \langle token list variable \rangle is not defined in this case and should not be relied upon. If the \langle comma list \rangle is non-empty, pops the top item from the \langle comma list \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle comma list \rangle. The \langle comma list \rangle is modified globally, while the \langle token list variable \rangle is assigned locally.

\clist_gpush:Nn \clist_gpush:

\clist_item:Nn \clist_item:cn \clist_item:nn

\clist_rand_item:N \clist_rand_item:c \clist_rand_item:n

\clist_show:N \clist_show:c

\new: 2012-05-14
\new: 2014-07-17
\new: 2016-12-06

8 Using a single item

\new: 2016-12-06

9 Viewing comma lists

\new: 2015-08-03

126
\clist_show:n \clist_show:n \{(tokens)\}
Displays the entries in the comma list in the terminal.

\clist_log:N \clist_log:N \langle comma list \rangle
Writes the entries in the \langle comma list \rangle in the log file. See also \clist_show:N which displays the result in the terminal.

\clist_log:n \clist_log:n \langle\{tokens\}\rangle
Writes the entries in the comma list in the log file. See also \clist_show:n which displays the result in the terminal.

10 Constant and scratch comma lists

\c_empty_clist
Constant that is always empty.

\l_tmpa_clist \l_tmpb_clist
Scratch comma lists for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_clist \g_tmpb_clist
Scratch comma lists for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
Part XVI
The \texttt{l3token} package
Token manipulation

This module deals with tokens. Now this is perhaps not the most precise description so let’s try with a better description: When programming in \TeX, it is often desirable to know just what a certain token is: is it a control sequence or something else. Similarly one often needs to know if a control sequence is expandable or not, a macro or a primitive, how many arguments it takes etc. Another thing of great importance (especially when it comes to document commands) is looking ahead in the token stream to see if a certain character is present and maybe even remove it or disregard other tokens while scanning.

This module provides functions for both and as such has two primary function categories: \texttt{token} for anything that deals with tokens and \texttt{peek} for looking ahead in the token stream.

Most functions we describe here can be used on control sequences, as those are tokens as well.

It is important to distinguish two aspects of a token: its “shape” (for lack of a better word), which affects the matching of delimited arguments and the comparison of token lists containing this token, and its “meaning”, which affects whether the token expands or what operation it performs. One can have tokens of different shapes with the same meaning, but not the converse.

For instance, \texttt{\if:w, \if_charcode:w, and \tex_if:D} are three names for the same internal operation of \TeX, namely the primitive testing the next two characters for equality of their character code. They have the same meaning hence behave identically in many situations. However, \TeX distinguishes them when searching for a delimited argument. Namely, the example function \texttt{\show_until_if:w} defined below takes everything until \texttt{\if:w} as an argument, despite the presence of other copies of \texttt{\if:w} under different names.

\begin{verbatim}
cs_new:Npn \show_until_if:w #1 \if:w { \tl_show:n {#1} }
\show_until_if:w \tex_if:D \if_charcode:w \if:w
\end{verbatim}

A list of all possible shapes and a list of all possible meanings are given in section \texttt{7}.

1 Creating character tokens

\begin{verbatim}
cchar_set_active_eq:NW \char_set_active_eq:Nc
\char_gset_active_eq:NW \char_gset_active_eq:Nc
\end{verbatim}

Updated: 2015-11-12

Sets the behaviour of the \texttt{⟨char⟩} in situations where it is active (category code 13) to be equivalent to that of the \texttt{⟨function⟩}. The category code of the \texttt{⟨char⟩} is \texttt{unchanged} by this process. The \texttt{⟨function⟩} may itself be an active character.

\begin{verbatim}
cchar_set_active_eq:nN \char_set_active_eq:nc
\char_gset_active_eq:nN \char_gset_active_eq:nc
\end{verbatim}

New: 2015-11-12

Sets the behaviour of the \texttt{⟨char⟩} which has character code as given by the \texttt{⟨integer expression⟩} in situations where it is active (category code 13) to be equivalent to that of the \texttt{⟨function⟩}. The category code of the \texttt{⟨char⟩} is \texttt{unchanged} by this process. The \texttt{⟨function⟩} may itself be an active character.
\char_generate:nn *(charcode) {catcode}*

Generates a character token of the given \textit{charcode} and \textit{catcode} (both of which may be integer expressions). The \textit{catcode} may be one of

- 1 (begin group)
- 2 (end group)
- 3 (math toggle)
- 4 (alignment)
- 6 (parameter)
- 7 (math superscript)
- 8 (math subscript)
- 10 (letter)
- 11 (other)
- 13 (active)

and other values raise an error. The \textit{charcode} may be any one valid for the engine in use. Active characters cannot be generated in older versions of \TeX{}. **\TeXhackers note:** Exactly two expansions are needed to produce the character.


Converting the \textit{char} to the equivalent case-changed character as detailed by the function name (see \texttt{\str_foldcase:n} and \texttt{\text_titlecase:n} for details of these terms). The case mapping is carried out with no context-dependence (\textit{cf.} \texttt{\text_uppercase:n}, \textit{etc.}) The \texttt{str} versions always generate "other" (category code 12) characters, whilst the standard versions generate characters with the category code of the \textit{char} (i.e. only the character code changes).

\c_catcode_other_space_tl

Token list containing one character with category code 12, ("other"), and character code 32 (space).
2 Manipulating and interrogating character tokens

\char_set_catcode_escape:N \char_set_catcode_letter:N \langle \text{character} \rangle
\char_set_catcode_group_begin:N \char_set_catcode_group_end:N
\char_set_catcode_math_toggle:N \char_set_catcode_alignment:N
\char_set_catcode_end_line:N \char_set_catcode_parameter:N
\char_set_catcode_math_superscript:N \char_set_catcode_math_subscript:N
\char_set_catcode_ignore:N \char_set_catcode_space:N
\char_set_catcode_letter:n {\langle \text{integer expression} \rangle}
\char_set_catcode_other:N \char_set_catcode_comment:n \char_set_catcode_invalid:n

Updated: 2015-11-11

Sets the category code of the \langle \text{character} \rangle to that indicated in the function name. Depending on the current category code of the \langle \text{token} \rangle the escape token may also be needed:

\char_set_catcode_other:n \% 

The assignment is local.

\char_set_catcode_escape:n \char_set_catcode_letter:n {\langle \text{integer expression} \rangle}
\char_set_catcode_group_begin:n \char_set_catcode_group_end:n
\char_set_catcode_math_toggle:n \char_set_catcode_alignment:n
\char_set_catcode_end_line:n \char_set_catcode_parameter:n
\char_set_catcode_math_superscript:n \char_set_catcode_math_subscript:n
\char_set_catcode_ignore:n \char_set_catcode_space:n
\char_set_catcode_letter:n \char_set_catcode_other:n
\char_set_catcode_active:n \char_set_catcode_comment:n \char_set_catcode_invalid:n

Updated: 2015-11-11

Sets the category code of the \langle \text{character} \rangle which has character code as given by the \langle \text{integer expression} \rangle. This version can be used to set up characters which cannot otherwise be given (\cf the N-type variants). The assignment is local.
These functions set the category code of the \textit{character} which has character code as given by the \textit{integer expression}. The first \textit{integer expression} is the character code and the second is the category code to apply. The setting applies within the current \TeX group. In general, the symbolic functions $\char_set_catcode_{\textit{type}}$ should be preferred, but there are cases where these lower-level functions may be useful.

Expands to the current category code of the \textit{character} with character code given by the \textit{integer expression}.

Displays the current category code of the \textit{character} with character code given by the \textit{integer expression} on the terminal.

Sets up the behaviour of the \textit{character} when found inside $\texttt{text_lowercase:n}$, such that \textit{character1} will be converted into \textit{character2}. The two \textit{characters} may be specified using an \textit{integer expression} for the character code concerned. This may include the \TeX \texttt{\char} method for converting a single character into its character code:

\begin{verbatim}
\char_set_lccode:nn { \char } { \char + 32 }
\char_set_lccode:nn { 60 } { 50 }
\end{verbatim}

The setting applies within the current \TeX group.

Expands to the current lower case code of the \textit{character} with character code given by the \textit{integer expression}.

Displays the current lower case code of the \textit{character} with character code given by the \textit{integer expression} on the terminal.

Sets up the behaviour of the \textit{character} when found inside $\texttt{text_uppercase:n}$, such that \textit{character1} will be converted into \textit{character2}. The two \textit{characters} may be specified using an \textit{integer expression} for the character code concerned. This may include the \TeX \texttt{\char} method for converting a single character into its character code:

\begin{verbatim}
\char_set_uccode:nn { \char } { \char - 32 }
\char_set_uccode:nn { 50 } { 60 }
\end{verbatim}

The setting applies within the current \TeX group.
\char_value_uccode:n \char_value_uccode:n \{\integer\ expression\}\}
\char_value_uccode:n \char_value_uccode:n \{\integer\ expression\}\}

Expands to the current upper case code of the \langle character\rangle with character code given by the \langle integer\ expression\rangle.

\char_show_uccode:n \char_show_uccode:n \{\integer\ expression\}\}
\char_show_uccode:n \char_show_uccode:n \{\integer\ expression\}\}

Displays the current upper case code of the \langle character\rangle with character code given by the \langle integer\ expression\rangle on the terminal.

\char_set_mathcode:nn
\char_set_mathcode:nn \{\intexpr\1\} \{\intexpr\2\}
\char_set_mathcode:nn
\char_set_mathcode:nn

This function sets up the math code of \langle character\rangle. The \langle character\rangle is specified as an \langle integer\ expression\rangle which will be used as the character code of the relevant character. The setting applies within the current \TeX\ group.

\char_value_mathcode:n \char_value_mathcode:n \{\integer\ expression\}\}
\char_value_mathcode:n \char_value_mathcode:n \{\integer\ expression\}\}

Expands to the current math code of the \langle character\rangle with character code given by the \langle integer\ expression\rangle.

\char_value_mathcode:n \char_value_mathcode:n \{\integer\ expression\}\}
\char_value_mathcode:n \char_value_mathcode:n \{\integer\ expression\}\}

Displays the current math code of the \langle character\rangle with character code given by the \langle integer\ expression\rangle on the terminal.

\char_set_sfcode:nn
\char_set_sfcode:nn \{\intexpr\1\} \{\intexpr\2\}
\char_set_sfcode:nn
\char_set_sfcode:nn

This function sets up the space factor for the \langle character\rangle. The \langle character\rangle is specified as an \langle integer\ expression\rangle which will be used as the character code of the relevant character. The setting applies within the current \TeX\ group.

\char_value_sfcode:n \char_value_sfcode:n \{\integer\ expression\}\}
\char_value_sfcode:n \char_value_sfcode:n \{\integer\ expression\}\}

Expands to the current space factor for the \langle character\rangle with character code given by the \langle integer\ expression\rangle.

\char_value_sfcode:n \char_value_sfcode:n \{\integer\ expression\}\}
\char_value_sfcode:n \char_value_sfcode:n \{\integer\ expression\}\}

Displays the current space factor for the \langle character\rangle with character code given by the \langle integer\ expression\rangle on the terminal.

\l_char_active_seq
\l_char_active_seq

Used to track which tokens may require special handling at the document level as they are (or have been at some point) of category \langle active\rangle (catcode 13). Each entry in the sequence consists of a single escaped token, for example \~. Active tokens should be added to the sequence when they are defined for general document use.

\l_char_special_seq
\l_char_special_seq

Updated: 2015-11-11

Updated: 2015-11-11

Used to track which tokens will require special handling when working with verbatim-like material at the document level as they are not of categories \langle letter\rangle (catcode 11) or \langle other\rangle (catcode 12). Each entry in the sequence consists of a single escaped token, for example \ for the backslash or \ for an opening brace. Escaped tokens should be added to the sequence when they are defined for general document use.
3 Generic tokens

These are implicit tokens which have the category code described by their name. They are used internally for test purposes but are also available to the programmer for other uses.

\c_group_begin_token
\c_group_end_token
\c_math_toggle_token
\c_alignment_token
\c_parameter_token
\c_math_superscript_token
\c_math_subscript_token
\c_space_token

These are implicit tokens which have the category code described by their name. They are used internally for test purposes and should not be used other than for category code tests.

\c_catcode_letter_token
\c_catcode_other_token

A token list containing an active token. This is used internally for test purposes and should not be used other than in appropriately-constructed category code tests.

4 Converting tokens

\token_to_meaning:N \token_to_meaning:c

Inserts the current meaning of the \langle token\rangle into the input stream as a series of characters of category code 12 (other). This is the primitive \TeX\ description of the \langle token\rangle, thus for example both functions defined by \cs_set_nopar:Npn and token list variables defined using \tl_new:N are described as macros.

\TeXhackers note: This is the \TeX\ primitive \meaning. The \langle token\rangle can thus be an explicit space tokens or an explicit begin-group or end-group character token ({ or } when normal \TeX\ category codes apply) even though these are not valid \textit{N}-type arguments.

\token_to_str:N \token_to_str:c

Converts the given \langle token\rangle into a series of characters with category code 12 (other). If the \langle token\rangle is a control sequence, this will start with the current escape character with category code 12 (the escape character is part of the \langle token\rangle). This function requires only a single expansion.

\TeXhackers note: \token_to_str:N is the \TeX\ primitive \string renamed. The \langle token\rangle can thus be an explicit space tokens or an explicit begin-group or end-group character token ({ or } when normal \TeX\ category codes apply) even though these are not valid \textit{N}-type arguments.
5 Token conditionals

\token_if_group_begin_p:N \token_if_group_begin:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of a begin group token \{ when normal \TeX\ category codes are in force). Note that an explicit begin group token cannot be tested in this way, as it is not a valid N-type argument.

\token_if_group_end_p:N \token_if_group_end:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of an end group token \} when normal \TeX\ category codes are in force). Note that an explicit end group token cannot be tested in this way, as it is not a valid N-type argument.

\token_if_math_toggle_p:N \token_if_math_toggle:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of a math shift token \$ when normal \TeX\ category codes are in force).

\token_if_alignment_p:N \token_if_alignment:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of an alignment token \& when normal \TeX\ category codes are in force).

\token_if_parameter_p:N \token_if_parameter:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of a macro parameter token \# when normal \TeX\ category codes are in force).

\token_if_math_superscript_p:N \token_if_math_superscript:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of a superscript token \^ when normal \TeX\ category codes are in force).

\token_if_math_subscript_p:N \token_if_math_subscript:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of a subscript token \_ when normal \TeX\ category codes are in force).

\token_if_space_p:N \token_if_space:N \{ (token) \} \{ (false code) \}

Tests if \(token\) has the category code of a space token. Note that an explicit space token with character code 32 cannot be tested in this way, as it is not a valid N-type argument.
\texttt{token\_if\_letter\_p:N} \texttt{*} \texttt{token\_if\_letter:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if \texttt{token} has the category code of a letter token.

\texttt{token\_if\_other\_p:N} \texttt{*} \texttt{token\_if\_other:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if \texttt{token} has the category code of an “other” token.

\texttt{token\_if\_active\_p:N} \texttt{*} \texttt{token\_if\_active:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if \texttt{token} has the category code of an active character.

\texttt{token\_if\_eq\_catcode\_p:NN} \texttt{*} \texttt{token\_if\_eq\_catcode:NN \{token\}_1 \{token\}_2}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the two \texttt{tokens} have the same category code.

\texttt{token\_if\_eq\_charcode\_p:NN} \texttt{*} \texttt{token\_if\_eq\_charcode:NN \{token\}_1 \{token\}_2}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the two \texttt{tokens} have the same character code.

\texttt{token\_if\_eq\_meaning\_p:NN} \texttt{*} \texttt{token\_if\_eq\_meaning:NN \{token\}_1 \{token\}_2}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the two \texttt{tokens} have the same meaning when expanded.

\texttt{token\_if\_macro\_p:N} \texttt{*} \texttt{token\_if\_macro:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the \texttt{token} is a \TeX{} macro.

\texttt{token\_if\_cs\_p:N} \texttt{*} \texttt{token\_if\_cs:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the \texttt{token} is a control sequence.

\texttt{token\_if\_expandable\_p:N} \texttt{*} \texttt{token\_if\_expandable:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the \texttt{token} is expandable. This test returns \texttt{false} for an undefined token.

\texttt{token\_if\_long\_macro\_p:N} \texttt{*} \texttt{token\_if\_long\_macro:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the \texttt{token} is a long macro.

\texttt{token\_if\_protected\_macro\_p:N} \texttt{*} \texttt{token\_if\_protected\_macro:N \{token\}}
\texttt{\{true code\}} \texttt{\{false code\}}

Tests if the \texttt{token} is a protected macro: for a macro which is both protected and long this returns \texttt{false}.

Updated: 2012-01-20
Tests if the \langle token \rangle is defined to be a chardef.

\textbf{TeXhackers note:} Boolean, boxes and small integer constants are implemented as \chardef s.

Tests if the \langle token \rangle is defined to be a mathchardef.

\textbf{TeXhackers note:} Constant integers may be implemented as integer registers, \chardef s, or \mathchardef s depending on their value.

Tests if the \langle token \rangle is defined to be a muskip register.

Tests if the \langle token \rangle is defined to be a skip register.
Tests if the \( \langle \text{token} \rangle \) is defined to be a toks register (not used by \LaTeX{}3).

Tests if the \( \langle \text{token} \rangle \) is an engine primitive.

### 6 Peeking ahead at the next token

There is often a need to look ahead at the next token in the input stream while leaving it in place. This is handled using the “peek” functions. The generic \texttt{\peek\_after:Nw} is provided along with a family of predefined tests for common cases. As peeking ahead does \textit{not} skip spaces the predefined tests include both a space-respecting and space-skipping version.

\begin{verbatim}
\peek\_after:Nw \texttt{\peek\_after:Nw \langle function \rangle \langle token \rangle}
\end{verbatim}

Locally sets the test variable \texttt{\_l\_peek\_token} equal to \( \langle \text{token} \rangle \) (as an implicit token, \textit{not} as a token list), and then expands the \( \langle \text{function} \rangle \). The \( \langle \text{token} \rangle \) remains in the input stream as the next item after the \( \langle \text{function} \rangle \). The \( \langle \text{token} \rangle \) here may be \texttt{\_\_\_ or } \texttt{\{ or \} } (assuming normal \TeX{} category codes), \textit{i.e.} it is not necessarily the next argument which would be grabbed by a normal function.

\begin{verbatim}
\peek\_gafter:Nw \texttt{\peek\_gafter:Nw \langle function \rangle \langle token \rangle}
\end{verbatim}

Globally sets the test variable \texttt{\_g\_peek\_token} equal to \( \langle \text{token} \rangle \) (as an implicit token, \textit{not} as a token list), and then expands the \( \langle \text{function} \rangle \). The \( \langle \text{token} \rangle \) remains in the input stream as the next item after the \( \langle \text{function} \rangle \). The \( \langle \text{token} \rangle \) here may be \texttt{\_\_\_ or } \texttt{\{ or \} } (assuming normal \TeX{} category codes), \textit{i.e.} it is not necessarily the next argument which would be grabbed by a normal function.

\texttt{\_l\_peek\_token}

Token set by \texttt{\peek\_after:Nw} and available for testing as described above.

\texttt{\_g\_peek\_token}

Token set by \texttt{\peek\_gafter:Nw} and available for testing as described above.

\begin{verbatim}
\peek\_catcode:NTF \texttt{\peek\_catcode:NTF \langle test token \rangle \{\langle true code \rangle\} \{\langle false code \rangle\}}
\end{verbatim}

Tests if the next \( \langle \text{token} \rangle \) in the input stream has the same category code as the \( \langle \text{test token} \rangle \) (as defined by the test \texttt{\token\_if\_eq\_catcode:NTF}). Spaces are respected by the test and the \( \langle \text{token} \rangle \) is left in the input stream after the \( \langle \text{true code} \rangle \) or \( \langle \text{false code} \rangle \) (as appropriate to the result of the test).
Tests if the next non-space \textit{(token)} in the input stream has the same category code as the \textit{(test token)} (as defined by the test \texttt{\token_if_eq_catcode:NNTF}). Explicit and implicit space tokens (with character code 32 and category code 10) are ignored and removed by the test and the \textit{(token)} is left in the input stream after the \texttt{(true code)} or \texttt{(false code)} (as appropriate to the result of the test).

Tests if the next \textit{(token)} in the input stream has the same category code as the \textit{(test token)} (as defined by the test \texttt{\token_if_eq_catcode:NNTF}). Spaces are respected by the test and the \textit{(token)} is removed from the input stream if the test is true. The function then places either the \texttt{(true code)} or \texttt{(false code)} in the input stream (as appropriate to the result of the test).

Tests if the next non-space \textit{(token)} in the input stream has the same character code as \textit{(test token)} (as defined by the test \texttt{\token_if_eq_charcode:NNTF}). Explicit and implicit space tokens (with character code 32 and category code 10) are ignored and removed by the test and the \textit{(token)} is left in the input stream after the \texttt{(true code)} or \texttt{(false code)} (as appropriate to the result of the test).

Tests if the next \textit{(token)} in the input stream has the same character code as the \textit{(test token)} (as defined by the test \texttt{\token_if_eq_charcode:NNTF}). Spaces are respected by the test and the \textit{(token)} is left in the input stream after the \texttt{(true code)} or \texttt{(false code)} (as appropriate to the result of the test).
Tests if the next non-space \{token\} in the input stream has the same character code as the \{test token\} (as defined by the test \texttt{token_if_eq_charcode:NNTF}). Explicit and implicit space tokens (with character code 32 and category code 10) are ignored and removed by the test and the \{token\} is removed from the input stream if the test is true. The function then places either the \{true code\} or \{false code\} in the input stream (as appropriate to the result of the test).

Tests if the next \{token\} in the input stream has the same meaning as the \{test token\} (as defined by the test \texttt{token_if_eq_meaning:NNTF}). Spaces are respected by the test and the \{token\} is left in the input stream after the \{true code\} or \{false code\} (as appropriate to the result of the test).

Tests if the next non-space \{token\} in the input stream has the same meaning as the \{test token\} (as defined by the test \texttt{token_if_eq_meaning:NNTF}). Explicit and implicit space tokens (with character code 32 and category code 10) are ignored and removed by the test and the \{token\} is removed from the input stream if the test is true. The function then places either the \{true code\} or \{false code\} in the input stream (as appropriate to the result of the test).

Tests if the next \{token\} in the input stream has the same meaning as the \{test token\} (as defined by the test \texttt{token_if_eq_meaning:NNTF}). Spaces are respected by the test and the \{token\} is removed from the input stream if the test is true. The function then places either the \{true code\} or \{false code\} in the input stream (as appropriate to the result of the test).
\peek N_type:TF \peek N_type:TF \{(true code)\} \{(false code)\}

Tests if the next \langle token \rangle in the input stream can be safely grabbed as an N-type argument. The test is \langle false \rangle if the next \langle token \rangle is either an explicit or implicit begin-group or end-group token (with any character code), or an explicit or implicit space character (with character code 32 and category code 10), or an outer token (never used in \TeX3) and \langle true \rangle in all other cases. Note that a \langle true \rangle result ensures that the next \langle token \rangle is a valid N-type argument. However, if the next \langle token \rangle is for instance \c_space_token, the test takes the \langle false \rangle branch, even though the next \langle token \rangle is in fact a valid N-type argument. The \langle token \rangle is left in the input stream after the \langle true code \rangle or \langle false code \rangle (as appropriate to the result of the test).

7 Description of all possible tokens

Let us end by reviewing every case that a given token can fall into. This section is quite technical and some details are only meant for completeness. We distinguish the meaning of the token, which controls the expansion of the token and its effect on \TeX's state, and its shape, which is used when comparing token lists such as for delimited arguments. Two tokens of the same shape must have the same meaning, but the converse does not hold.

A token has one of the following shapes.

- A control sequence, characterized by the sequence of characters that constitute its name: for instance, \use:n is a five-letter control sequence.
- An active character token, characterized by its character code (between 0 and 1114111 for Lua\TeX and XT\Ex and less for other engines) and category code 13.
- A character token, characterized by its character code and category code (one of 1, 2, 3, 4, 6, 7, 8, 10, 11 or 12 whose meaning is described below).\(^7\)

There are also a few internal tokens. The following list may be incomplete in some engines.

- Expanding \texttt{\the\font} results in a token that looks identical to the command that was used to select the current font (such as \texttt{\tenrm}) but it differs from it in shape.
- A “frozen” \relax, which differs from the primitive in shape (but has the same meaning), is inserted when the closing \texttt{\fi} of a conditional is encountered before the conditional is evaluated.
- Expanding \texttt{\noexpand \{token\}} (when the \langle token \rangle is expandable) results in an internal token, displayed (temporarily) as \texttt{\notexpanded: \{token\}}, whose shape coincides with the \langle token \rangle and whose meaning differs from \texttt{\relax}.
- An \texttt{\outer endtemplate:} can be encountered when peeking ahead at the next token; this expands to another internal token, \texttt{end of alignment template}.
- Tricky programming might access a frozen \texttt{\endwrite}.

\(^7\)In Lua\TeX, there is also the case of “bytes”, which behave as character tokens of category code 12 (other) and character code between 1114112 and 1114366. They are used to output individual bytes to files, rather than UTF-8.
• Some frozen tokens can only be accessed in interactive sessions: \cr, \right, \endgroup, \fi, \inaccessible.

The meaning of a (non-active) character token is fixed by its category code (and character code) and cannot be changed. We call these tokens explicit character tokens.

Category codes that a character token can have are listed below by giving a sample output of the \TeX{} primitive \meaning, together with their \LaTEX{} names and most common example:

1. begin-group character (\texttt{group\_begin}, often \{),
2. end-group character (\texttt{group\_end}, often \}),
3. math shift character (\texttt{math\_toggle}, often \$),
4. alignment tab character (\texttt{alignment}, often &),
5. macro parameter character (\texttt{parameter}, often \#),
6. superscript character (\texttt{math\_superscript}, often ^),
7. subscript character (\texttt{math\_subscript}, often _),
8. blank space (\texttt{space}, often character code 32),
9. the letter (\texttt{letter}, such as A),
10. the character (\texttt{other}, such as 0).

Category code 13 (active) is discussed below. Input characters can also have several other category codes which do not lead to character tokens for later processing: 0 (\texttt{escape}), 5 (\texttt{end\_line}), 9 (\texttt{ignore}), 14 (\texttt{comment}), and 15 (\texttt{invalid}).

The meaning of a control sequence or active character can be identical to that of any character token listed above (with any character code), and we call such tokens implicit character tokens. The meaning is otherwise in the following list:

• a macro, used in \LaTEX{} for most functions and some variables (\texttt{tl}, \texttt{fp}, \texttt{seq}, ...),
• a primitive such as \texttt{\def} or \texttt{\topmark}, used in \LaTEX{} for some functions,
• a register such as \texttt{\count123}, used in \LaTEX{} for the implementation of some variables (\texttt{int}, \texttt{dim}, ...),
• a constant integer such as \texttt{\char"56} or \texttt{\mathchar"121},
• a font selection command,
• undefined.

Macros be \texttt{\protected} or not, \texttt{\long} or not (the opposite of what \LaTEX{} calls \texttt{nopar}), and \texttt{\outer} or not (unused in \LaTEX{}). Their \meaning takes the form

\langle properties \rangle \texttt{macro:} \langle parameters \rangle \Rightarrow \langle replacement \rangle
where \(\langle\text{properties}\rangle\) is among \texttt{\protect\long\outer}, \(\langle\text{parameters}\rangle\) describes parameters that the macro expects, such as \#1\#2\#3, and \(\langle\text{replacement}\rangle\) describes how the parameters are manipulated, such as \#2/#1/#3.

Now is perhaps a good time to mention some subtleties relating to tokens with category code 10 (space). Any input character with this category code (normally, space and tab characters) becomes a normal space, with character code 32 and category code 10.

When a macro takes an undelimited argument, explicit space characters (with character code 32 and category code 10) are ignored. If the following token is an explicit character token with category code 1 (begin-group) and an arbitrary character code, then \TeX{} scans ahead to obtain an equal number of explicit character tokens with category code 1 (begin-group) and 2 (end-group), and the resulting list of tokens (with outer braces removed) becomes the argument. Otherwise, a single token is taken as the argument for the macro: we call such single tokens “N-type”, as they are suitable to be used as an argument for a function with the signature \(:\text{N}\).
Part XVII

The \l3prop package

Property lists

\LaTeX3 implements a “property list” data type, which contain an unordered list of entries each of which consists of a \langle \textit{key} \rangle and an associated \langle \textit{value} \rangle. The \langle \textit{key} \rangle and \langle \textit{value} \rangle may both be any \langle \textit{balanced text} \rangle. It is possible to map functions to property lists such that the function is applied to every key–value pair within the list.

Each entry in a property list must have a unique \langle \textit{key} \rangle: if an entry is added to a property list which already contains the \langle \textit{key} \rangle then the new entry overwrites the existing one. The \langle \textit{keys} \rangle are compared on a string basis, using the same method as \texttt{\str_if_eq:nn}.

Property lists are intended for storing key-based information for use within code. This is in contrast to key–value lists, which are a form of \texttt{input} parsed by the \texttt{keys} module.

1 Creating and initialising property lists

\begin{verbatim}
\prop_new:N \prop_new:c
\prop_clear:N \prop_clear:c \prop_gclear:N \prop_gclear:c
\prop_set_eq:NN \prop_set_eq:cn \prop_gset_eq:NN \prop_gset_eq:cn
\prop_set_from_keyval:Nn \prop_set_from_keyval:cn \prop_gset_from_keyval:NN \prop_gset_from_keyval:cn
\end{verbatim}

\texttt{\prop_new:N \textit{property list}}

Creates a new \langle \textit{property list} \rangle or raises an error if the name is already taken. The declaration is global. The \langle \textit{property list} \rangle initially contains no entries.

\texttt{\prop_clear:N \textit{property list}}

Clears all entries from the \langle \textit{property list} \rangle.

\texttt{\prop_set_eq:NN \textit{property list}1 \textit{property list}2}

Sets the content of \langle \textit{property list}1 \rangle equal to that of \langle \textit{property list}2 \rangle.

\texttt{\prop_set_from_keyval:Nn \prop_set_from_keyval:cn \prop_gset_from_keyval:NN \prop_gset_from_keyval:cn}

New: 2017-11-28

Updated: 2019-08-25

 Sets \langle \textit{prop var} \rangle to contain key–value pairs given in the second argument. If duplicate keys appear only one of the values is kept.
\prop_const_from_keyval:Nn
\prop_const_from_keyval:cn

New: 2017-11-28
Updated: 2019-08-25

Creates a new constant \langle \text{prop var} \rangle or raises an error if the name is already taken. The \langle \text{prop var} \rangle is set globally to contain key–value pairs given in the second argument. If duplicate keys appear only one of the values is kept.

2 Adding entries to property lists

\prop_put:Nnn
\prop_put:NnN
\prop_gput:Nnn
\prop_gput:NnN

Updated: 2012-07-09

Adds an entry to the \langle property list \rangle which may be accessed using the \langle key \rangle and which has \langle value \rangle. Both the \langle key \rangle and \langle value \rangle may contain any \langle balanced text \rangle. The \langle key \rangle is stored after processing with \tl_to_str:n, meaning that category codes are ignored. If the \langle key \rangle is already present in the \langle property list \rangle, the existing entry is overwritten by the new \langle value \rangle.

\prop_put_if_new:Nnn
\prop_put_if_new:NnN
\prop_gput_if_new:Nnn
\prop_gput_if_new:NnN

If the \langle key \rangle is present in the \langle property list \rangle then no action is taken. If the \langle key \rangle is not present in the \langle property list \rangle then a new entry is added. Both the \langle key \rangle and \langle value \rangle may contain any \langle balanced text \rangle. The \langle key \rangle is stored after processing with \tl_to_str:n, meaning that category codes are ignored.

3 Recovering values from property lists

\prop_get:NnN
\prop_get:NnN

Updated: 2011-08-28

Recovers the \langle value \rangle stored with \langle key \rangle from the \langle property list \rangle, and places this in the \langle token list variable \rangle. If the \langle key \rangle is not found in the \langle property list \rangle then the \langle token list variable \rangle is set to the special marker \q_no_value. The \langle token list variable \rangle is set within the current \TeX\ group. See also \prop_get:NnNTF.

\prop_pop:NnN
\prop_pop:NnN

Updated: 2011-08-18

Recovers the \langle value \rangle stored with \langle key \rangle from the \langle property list \rangle, and places this in the \langle token list variable \rangle. If the \langle key \rangle is not found in the \langle property list \rangle then the \langle token list variable \rangle is set to the special marker \q_no_value. The \langle key \rangle and \langle value \rangle are then deleted from the property list. Both assignments are local. See also \prop_pop:NnNTF.
\prop_gpop:NnN
\prop_gpop:{NoN}[cn][coN]

Updated: 2011-08-18

\prop_gpop:NnN (property list) {⟨key⟩} ⟨tl var⟩
Recovers the ⟨value⟩ stored with ⟨key⟩ from the ⟨property list⟩, and places this in the ⟨token list variable⟩. If the ⟨key⟩ is not found in the ⟨property list⟩ then the ⟨token list variable⟩ is set to the special marker \q_{no\_value}. The ⟨key⟩ and ⟨value⟩ are then deleted from the property list. The ⟨property list⟩ is modified globally, while the assignment of the ⟨token list variable⟩ is local. See also \prop_gpop:NnNTF.

\prop_item:Nn  *
\prop_item:cn  *
New: 2014-07-17

\prop_item:Nn (property list) {⟨key⟩}
Expands to the ⟨value⟩ corresponding to the ⟨key⟩ in the ⟨property list⟩. If the ⟨key⟩ is missing, this has an empty expansion.

\TeXhackers note: This function is slower than the non-expandable analogue \prop_get:Nn. The result is returned within the \unexpanded primitive \exp_not:n, which means that the ⟨value⟩ does not expand further when appearing in an x-type argument expansion.

\prop_count:N  *
\prop_count:c  *

Leaves the number of key–value pairs in the ⟨property list⟩ in the input stream as an ⟨integer denotation⟩.

4 Modifying property lists

\prop_remove:Nn (property list) {⟨key⟩}
Removes the entry listed under ⟨key⟩ from the ⟨property list⟩. If the ⟨key⟩ is not found in the ⟨property list⟩ no change occurs, i.e. there is no need to test for the existence of a key before deleting it.

5 Property list conditionals

\prop_if_exist_p:N  *
\prop_if_exist:p:c  *
\prop_if_exist:NTF  *
\prop_if_exist:cTF  *
New: 2012-03-03

\prop_if_exist_p:N (property list)
\prop_if_exist:NTF (property list) {⟨true code⟩} {⟨false code⟩}
Tests whether the ⟨property list⟩ is currently defined. This does not check that the ⟨property list⟩ really is a property list variable.

\prop_if_empty_p:N  *
\prop_if_empty:p:c  *
\prop_if_empty:NTF  *
\prop_if_empty:cTF  *

\prop_if_empty_p:N (property list)
\prop_if_empty:NTF (property list) {⟨true code⟩} {⟨false code⟩}
Tests if the ⟨property list⟩ is empty (containing no entries).
Tests if the \textit{key} is present in the \textit{property list}, making the comparison using the method described by $\texttt{\backslash str\_if\_eq:nnTF}$.

\textbf{\TeX{}hackers note:} This function iterates through every key–value pair in the \textit{property list} and is therefore slower than using the non-expandable $\texttt{\backslash prop\_get:NNTF}$.

\section{Recovering values from property lists with branching}

The functions in this section combine tests for the presence of a key in a property list with recovery of the associated valued. This makes them useful for cases where different cases follow dependent on the presence or absence of a key in a property list. They offer increased readability and performance over separate testing and recovery phases.

If the \textit{key} is not present in the \textit{property list}, leaves the \textit{false code} in the input stream. The value of the \textit{token list variable} is not defined in this case and should not be relied upon. If the \textit{key} is present in the \textit{property list}, stores the corresponding \textit{value} in the \textit{token list variable} without removing it from the \textit{property list}, then leaves the \textit{true code} in the input stream. The \textit{token list variable} is assigned locally.

If the \textit{key} is not present in the \textit{property list}, leaves the \textit{false code} in the input stream. The value of the \textit{token list variable} is not defined in this case and should not be relied upon. If the \textit{key} is present in the \textit{property list}, pops the corresponding \textit{value} in the \textit{token list variable}, i.e. removes the item from the \textit{property list}. Both the \textit{property list} and the \textit{token list variable} are assigned locally.

If the \textit{key} is not present in the \textit{property list}, leaves the \textit{false code} in the input stream. The value of the \textit{token list variable} is not defined in this case and should not be relied upon. If the \textit{key} is present in the \textit{property list}, pops the corresponding \textit{value} in the \textit{token list variable}, i.e. removes the item from the \textit{property list}. The \textit{property list} is modified globally, while the \textit{token list variable} is assigned locally.
7 Mapping to property lists

All mappings are done at the current group level, \textit{i.e.} any local assignments made by the \langle function \rangle or \langle code \rangle discussed below remain in effect after the loop.

\begin{verbatim}
\prop_map_function:NN \langle property list \rangle \langle function \rangle
\end{verbatim}

Applies \langle function \rangle to every \langle entry \rangle stored in the \langle property list \rangle. The \langle function \rangle receives two arguments for each iteration: the \langle key \rangle and associated \langle value \rangle. The order in which \langle entries \rangle are returned is not defined and should not be relied upon. To pass further arguments to the \langle function \rangle, see \prop_map_tokens:Nn.

\begin{verbatim}
\prop_map_function:cN
\end{verbatim}

\begin{verbatim}
\prop_map_inline:Nn \langle property list \rangle \{ \langle inline function \rangle \}
\end{verbatim}

Applies \langle inline function \rangle to every \langle entry \rangle stored within the \langle property list \rangle. The \langle inline function \rangle should consist of code which receives the \langle key \rangle as \#1 and the \langle value \rangle as \#2. The order in which \langle entries \rangle are returned is not defined and should not be relied upon.

\begin{verbatim}
\prop_map_inline:cn
\end{verbatim}

\begin{verbatim}
\prop_map_tokens:Nn \langle property list \rangle \{ \langle code \rangle \}
\end{verbatim}

Analogue of \prop_map_function:NN which maps several tokens instead of a single function. The \langle code \rangle receives each key–value pair in the \langle property list \rangle as two trailing brace groups. For instance,

\begin{verbatim}
\prop_map_tokens:Nn \l_my_prop \{ \str_if_eq:nnT \{ mykey \} \}
\end{verbatim}

expands to the value corresponding to \texttt{mykey}: for each pair in \l_my_prop the function \str_if_eq:nnT receives \texttt{mykey}, the \langle key \rangle and the \langle value \rangle as its three arguments. For that specific task, \prop_item:Nn is faster.

\begin{verbatim}
\prop_map_break:
\end{verbatim}

Used to terminate a \prop_map\ldots function before all entries in the \langle property list \rangle have been processed. This normally takes place within a conditional statement, for example

\begin{verbatim}
\prop_map_inline:Nn \l_my_prop \{ \str_if_eq:nnTF \{ #1 \} { bingo } \{ \prop_map_break: \} \{ % Do something useful \} \}
\end{verbatim}

Use outside of a \prop_map\ldots scenario leads to low level \TeX{} errors.

\textbf{\TeX{} hackers note:} When the mapping is broken, additional tokens may be inserted before further items are taken from the input stream. This depends on the design of the mapping function.
\prop_map_break:n \(\{\text{code}\}\)

Used to terminate a \prop_map_... function before all entries in the \textit{property list} have been processed, inserting the \textit{(code)} after the mapping has ended. This normally takes place within a conditional statement, for example

\prop_map_inline:Nn \l_my_prop
{\str_if_eq:nnTF { #1 } { bingo } { \prop_map_break:n { <code> } } \%
% Do something useful
}

Use outside of a \prop_map_... scenario leads to low level \TeX{} errors.

\TeX{}hackers note: When the mapping is broken, additional tokens may be inserted before the \textit{(code)} is inserted into the input stream. This depends on the design of the mapping function.

8 Viewing property lists

\prop_show:N \prop_show:c

Displays the entries in the \textit{property list} in the terminal.

\prop_log:N \prop_log:c

Writes the entries in the \textit{property list} in the log file.

9 Scratch property lists

\l_tmpa_prop \l_tmpb_prop

Scratch property lists for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX{}3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_prop \g_tmpb_prop

Scratch property lists for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX{}3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
10 Constants

\texttt{\textbackslash c\_empty\_prop} A permanently-empty property list used for internal comparisons.
Part XVIII
The \texttt{l3msg} package

Messages need to be passed to the user by modules, either when errors occur or to indicate how the code is proceeding. The \texttt{l3msg} module provides a consistent method for doing this (as opposed to writing directly to the terminal or log).

The system used by \texttt{l3msg} to create messages divides the process into two distinct parts. Named messages are created in the first part of the process; at this stage, no decision is made about the type of output that the message will produce. The second part of the process is actually producing a message. At this stage a choice of message \textit{class} has to be made, for example \texttt{error}, \texttt{warning} or \texttt{info}.

By separating out the creation and use of messages, several benefits are available. First, the messages can be altered later without needing details of where they are used in the code. This makes it possible to alter the language used, the detail level and so on. Secondly, the output which results from a given message can be altered. This can be done on a message class, module or message name basis. In this way, message behaviour can be altered and messages can be entirely suppressed.

1 Creating new messages

All messages have to be created before they can be used. The text of messages is automatically wrapped to the length available in the console. As a result, formatting is only needed where it helps to show meaning. In particular, \textbackslash{} may be used to force a new line and \textbackslash{} forces an explicit space. Additionally, \textbackslash{}l, \textbackslash{}#, \textbackslash{}%, \textbackslash{}# and \textbackslash{}- can be used to produce the corresponding character.

Messages may be subdivided \textit{by one level} using the / character. This is used within the message filtering system to allow for example the \LaTeX{} kernel messages to belong to the module \LaTeX{} while still being filterable at a more granular level. Thus for example

\begin{verbatim}
\msg_new:nnnn { mymodule } { submodule / message } ...
\end{verbatim}

will allow to filter out specifically messages from the \texttt{submodule}.

\begin{verbatim}
\msg_new:nnnn \msg_new:nnn \msg_set:nnnn \msg_set:nnn \msg_gset:nnnn \msg_gset:nnn
\end{verbatim}

\begin{verbatim}
\msg_new:nnnn \{(module)\} \{(message)\} \{(text)\} \{(more text)\}
\end{verbatim}

Creates a \langle message\rangle for a given \langle module\rangle. The message is defined to first give \langle text\rangle and then \langle more text\rangle if the user requests it. If no \langle more text\rangle is available then a standard text is given instead. Within \langle text\rangle and \langle more text\rangle four parameters (#1 to #4) can be used: these will be supplied at the time the message is used. An error is raised if the \langle message\rangle already exists.

\begin{verbatim}
\msg_set:nnnn \msg_gset:nnnn
\end{verbatim}

\begin{verbatim}
\msg_set:nnnn \{(module)\} \{(message)\} \{(text)\} \{(more text)\}
\end{verbatim}

Sets up the text for a \langle message\rangle for a given \langle module\rangle. The message is defined to first give \langle text\rangle and then \langle more text\rangle if the user requests it. If no \langle more text\rangle is available then a standard text is given instead. Within \langle text\rangle and \langle more text\rangle four parameters (#1 to #4) can be used: these will be supplied at the time the message is used.
Tests whether the \textit{message} for the \textit{module} is currently defined.

\section{Contextual information for messages}

\texttt{\msg_line_context:}\quad Prints the current line number when a message is given, and thus suitable for giving context to messages. The number itself is proceeded by the text \texttt{on line}.

\texttt{\msg_line_number:}\quad Prints the current line number when a message is given.

\texttt{\msg_fatal_text:n}\quad Produces the standard text

\textit{Fatal Package \textit{\textit{module}}} Error

This function can be redefined to alter the language in which the message is given, using \texttt{#1} as the name of the \textit{\textit{module}} to be included.

\texttt{\msg_critical_text:n}\quad Produces the standard text

\textit{Critical Package \textit{\textit{module}}} Error

This function can be redefined to alter the language in which the message is given, using \texttt{#1} as the name of the \textit{\textit{module}} to be included.

\texttt{\msg_error_text:n}\quad Produces the standard text

\textit{Package \textit{\textit{module}}} Error

This function can be redefined to alter the language in which the message is given, using \texttt{#1} as the name of the \textit{\textit{module}} to be included.

\texttt{\msg_warning_text:n}\quad Produces the standard text

\textit{Package \textit{\textit{module}}} Warning

This function can be redefined to alter the language in which the message is given, using \texttt{#1} as the name of the \textit{\textit{module}} to be included. The \textit{\textit{\textit{type}}} of \textit{\textit{\textit{module}}} may be adjusted: Package is the standard outcome: see \texttt{\msg_module_type:n}. 

151
\msg_info:n * \msg_info:n \{\langle module\rangle\}

 Produces the standard text:

 **Package** \langle module\rangle **Info**

 This function can be redefined to alter the language in which the message is given, using
 #1 as the name of the \langle module\rangle to be included. The \langle type\rangle of \langle module\rangle may be adjusted:
 **Package** is the standard outcome: see \msg_module_type:n.

\msg_module_name:n * \msg_module_name:n \{\langle module\rangle\}

 New: 2018-10-10

 Expands to the public name of the \langle module\rangle as defined by \g_msg_module_name_prop
 (or otherwise leaves the \langle module\rangle unchanged).

\msg_module_type:n * \msg_module_type:n \{\langle module\rangle\}

 New: 2018-10-10

 Expands to the description which applies to the \langle module\rangle, for example a **Package** or
 **Class**. The information here is defined in \g_msg_module_type_prop, and will default
 to **Package** if an entry is not present.

\msg_see_documentation_text:n * \msg_see_documentation_text:n \{\langle module\rangle\}

 Updated: 2018-09-30

 Produces the standard text

 **See the** \langle module\rangle **documentation for further information.**

 This function can be redefined to alter the language in which the message is given, using
 #1 as the name of the \langle module\rangle to be included. The name of the \langle module\rangle may be altered
 by use of \g_msg_module_documentation_prop

\g_msg_module_name_prop

 New: 2018-10-10

 Provides a mapping between the module name used for messages, and that for documenta-
 tion. For example, \LaTeX3 core messages are stored in the reserved **LaTeX** tree, but are
 printed as **LaTeX3**.

\g_msg_module_type_prop

 New: 2018-10-10

 Provides a mapping between the module name used for messages, and that type of
 module. For example, for \LaTeX3 core messages, an empty entry is set here meaning that
 they are not described using the standard **Package** text.

### 3 Issuing messages

Messages behave differently depending on the message class. In all cases, the message
 may be issued supplying 0 to 4 arguments. If the number of arguments supplied here does
 not match the number in the definition of the message, extra arguments are ignored, or
 empty arguments added (of course the sense of the message may be impaired). The four
 arguments are converted to strings before being added to the message text: the x-type
 variants should be used to expand material.
\msg_fatal:nnnnnn \{module\} \{message\} \{\arg one\} \{\arg two\} \{\arg three\} \{\arg four\}

Issues \langle module \rangle \text{ error } \langle message \rangle, \text{ passing } \langle \arg one \rangle \text{ to } \langle \arg four \rangle \text{ to the text-creating functions. After issuing a fatal error the } \LaTeX \text{ run halts. No PDF file will be produced in this case (DVI mode runs may produce a truncated DVI file).}

\msg_critical:nnnnnn \{module\} \{message\} \{\arg one\} \{\arg two\} \{\arg three\} \{\arg four\}

Issues \langle module \rangle \text{ error } \langle message \rangle, \text{ passing } \langle \arg one \rangle \text{ to } \langle \arg four \rangle \text{ to the text-creating functions. After issuing a critical error, } \LaTeX \text{ stops reading the current input file. This may halt the } \LaTeX \text{ run (if the current file is the main file) or may abort reading a sub-file.}

\textbf{\LaTeXhackers note:} The \LaTeX \texttt{\endinput} primitive is used to exit the file. In particular, the rest of the current line remains in the input stream.

\msg_error:nnnnnn \{module\} \{message\} \{\arg one\} \{\arg two\} \{\arg three\} \{\arg four\}

Issues \langle module \rangle \text{ error } \langle message \rangle, \text{ passing } \langle \arg one \rangle \text{ to } \langle \arg four \rangle \text{ to the text-creating functions. The error interrupts processing and issues the text at the terminal. After user input, the run continues.}

\msg_warning:nnnnn \{module\} \{message\} \{\arg one\} \{\arg two\} \{\arg three\} \{\arg four\}

Issues \langle module \rangle \text{ warning } \langle message \rangle, \text{ passing } \langle \arg one \rangle \text{ to } \langle \arg four \rangle \text{ to the text-creating functions. The warning text is added to the log file and the terminal, but the } \LaTeX \text{ run is not interrupted.}
Issues (module) information (message), passing (arg one) to (arg four) to the text-creating functions. The information text is added to the log file.

Issues (module) information (message), passing (arg one) to (arg four) to the text-creating functions. The information text is added to the log file: the output is briefer than \msg_info:nnnnnn.

Does nothing: used as a message class to prevent any output at all (see the discussion of message redirection).

4 Redirecting messages

Each message has a “name”, which can be used to alter the behaviour of the message when it is given. Thus we might have

\msg_new:nnnn { module } { my-message } { Some-text } { Some-more-text } to define a message, with

\msg_error:nn { module } { my-message }

when it is used. With no filtering, this raises an error. However, we could alter the behaviour with

\msg_redirect_class:nn { error } { warning }

to turn all errors into warnings, or with

\msg_redirect_module:nnn { module } { error } { warning }


to alter only messages from that module, or even

\msg_redirect_name:nnn { module } { my-message } { warning }

to target just one message. Redirection applies first to individual messages, then to messages from one module and finally to messages of one class. Thus it is possible to select out an individual message for special treatment even if the entire class is already redirected.

Multiple redirections are possible. Redirections can be cancelled by providing an empty argument for the target class. Redirection to a missing class raises an error immediately. Infinite loops are prevented by eliminating the redirection starting from the target of the redirection that caused the loop to appear. Namely, if redirections are requested as \( A \to B, B \to C \) and \( C \to A \) in this order, then the \( A \to B \) redirection is cancelled.

\msg_redirect_class:nn

\msg_redirect_class:nnn \{⟨class one⟩\} \{⟨class two⟩\}

Changes the behaviour of messages of \( ⟨\text{class one}⟩ \) so that they are processed using the code for those of \( ⟨\text{class two}⟩ \).

\msg_redirect_module:nnn

\msg_redirect_module:nnn \{⟨module⟩\} \{⟨class one⟩\} \{⟨class two⟩\}

Redirects message of \( ⟨\text{class one}⟩ \) for \( ⟨\text{module}⟩ \) to act as though they were from \( ⟨\text{class two}⟩ \). Messages of \( ⟨\text{class one}⟩ \) from sources other than \( ⟨\text{module}⟩ \) are not affected by this redirection. This function can be used to make some messages “silent” by default. For example, all of the \( \text{warning} \) messages of \( ⟨\text{module}⟩ \) could be turned off with:

\msg_redirect_module:nnn { module } { warning } { none }

\msg_redirect_name:nnn

\msg_redirect_name:nnn \{⟨module⟩\} \{⟨message⟩\} \{⟨class⟩\}

Redirects a specific \( ⟨\text{message}⟩ \) from a specific \( ⟨\text{module}⟩ \) to act as a member of \( ⟨\text{class}⟩ \) of messages. No further redirection is performed. This function can be used to make a selected message “silent” without changing global parameters:

\msg_redirect_name:nnn { module } { annoying-message } { none }
Part XIX

The \texttt{l3}file package

File and I/O operations

This module provides functions for working with external files. Some of these functions apply to an entire file, and have prefix \texttt{\file...}, while others are used to work with files on a line by line basis and have prefix \texttt{\ior...} (reading) or \texttt{\iow...} (writing).

It is important to remember that when reading external files \TeX{} attempts to locate them using both the operating system path and entries in the \TeX{} file database (most \TeX{} systems use such a database). Thus the “current path” for \TeX{} is somewhat broader than that for other programs.

For functions which expect a \texttt{(file name)} argument, this argument may contain both literal items and expandable content, which should on full expansion be the desired file name. Active characters (as declared in \texttt{\l_char_active_seq}) are not expanded, allowing the direct use of these in file names. Quote tokens (") are not permitted in file names as they are reserved for internal use by some \TeX{} primitives.

Spaces are trimmed at the beginning and end of the file name: this reflects the fact that some file systems do not allow or interact unpredictably with spaces in these positions. When no extension is given, this will trim spaces from the start of the name only.

1 Input–output stream management

As \TeX{} engines have a limited number of input and output streams, direct use of the streams by the programmer is not supported in \LaTeX3. Instead, an internal pool of streams is maintained, and these are allocated and deallocated as needed by other modules. As a result, the programmer should close streams when they are no longer needed, to release them for other processes.

Note that I/O operations are global: streams should all be declared with global names and treated accordingly.

\begin{code}
\ior_new:N \ior_new:c
\iow_new:N \iow_new:c
\ior_open:Nn \ior_open:cn
\iow_open:Nn \iow_open:cn
\end{code}

Globally reserves the name of the \texttt{(stream)}, either for reading or for writing as appropriate. The \texttt{(stream)} is not opened until the appropriate \texttt{\..._open:Nn} function is used. Attempting to use a \texttt{(stream)} which has not been opened is an error, and the \texttt{(stream)} will behave as the corresponding \texttt{\c_term...}.

Opens \texttt{(file name)} for reading using \texttt{(stream)} as the control sequence for file access. If the \texttt{(stream)} was already open it is closed before the new operation begins. The \texttt{(stream)} is available for access immediately and will remain allocated to \texttt{(file name)} until a \texttt{\ior_close:N} instruction is given or the \TeX{} run ends. If the file is not found, an error is raised.
\ior_open:NnTF \ior_open:cnTF

Opens \texttt{file name} for reading using \texttt{stream} as the control sequence for file access. If the \texttt{stream} was already open it is closed before the new operation begins. The \texttt{stream} is available for access immediately and will remain allocated to \texttt{file name} until a \ior_close:N instruction is given or the \TeX run ends. The \texttt{true code} is then inserted into the input stream. If the file is not found, no error is raised and the \texttt{false code} is inserted into the input stream.

\ior_close:N \ior_close:c \ior_close:N \ior_close:c

Closes the \texttt{stream}. Streams should always be closed when they are finished with as this ensures that they remain available to other programmers.

\ior_show_list: \ior_log_list: \ior_show_list: \ior_log_list:

Display (to the terminal or log file) a list of the file names associated with each open (read or write) stream. This is intended for tracking down problems.

1.1 Reading from files

Reading from files and reading from the terminal are separate processes in \texttt{expl3}. The functions \ior_get:NN and \ior_str_get:NN, and their branching equivalents, are designed to work with files.
Function that reads one or more lines (until an equal number of left and right braces are found) from the file input \langle stream \rangle and stores the result locally in the \langle token list variable \rangle variable. The material read from the \langle stream \rangle is tokenized by \TeX according to the category codes and \endlinechar in force when the function is used. Assuming normal settings, any lines which do not end in a comment character % have the line ending converted to a space, so for example input

\begin{verbatim}
a b c
\end{verbatim}

results in a token list a\,b\,c. Any blank line is converted to the token \par. Therefore, blank lines can be skipped by using a test such as

\begin{verbatim}
\ior_get:NN \l_my_stream \l_tmpa_tl
\tl_set:Nn \l_tmpb_tl { \par }
\tl_if_eq:NNF \l_tmpa_tl \l_tmpb_tl
...
\end{verbatim}

Also notice that if multiple lines are read to match braces then the resulting token list can contain \par tokens. In the non-branching version, where the \langle stream \rangle is not open the \langle tl var \rangle is set to \q_no_value.

\TeXhackers note: This protected macro is a wrapper around the \TeX primitive \read. Regardless of settings, \TeX replaces trailing space and tab characters (character codes 32 and 9) in each line by an end-of-line character (character code \endlinechar, omitted if \endlinechar is negative or too large) before turning characters into tokens according to current category codes. With default settings, spaces appearing at the beginning of lines are also ignored.

Function that reads one line from the file input \langle stream \rangle and stores the result locally in the \langle token list variable \rangle variable. The material is read from the \langle stream \rangle as a series of tokens with category code 12 (other), with the exception of space characters which are given category code 10 (space). Multiple whitespace characters are retained by this process. It always only reads one line and any blank lines in the input result in the \langle token list variable \rangle being empty. Unlike \ior_get:NN, line ends do not receive any special treatment. Thus input

\begin{verbatim}
a b c
\end{verbatim}

results in a token list a\,b\,c with the letters a, b, and c having category code 12. In the non-branching version, where the \langle stream \rangle is not open the \langle tl var \rangle is set to \q_no_value.

\TeXhackers note: This protected macro is a wrapper around the \epsilon-\TeX primitive \readline. Regardless of settings, \TeX removes trailing space and tab characters (character codes 32 and 9). However, the end-line character normally added by this primitive is not included in the result of \ior_str_get:NN.

All mappings are done at the current group level, \ie any local assignments made by the \langle function \rangle or \langle code \rangle discussed below remain in effect after the loop.
\ior_map_inline:Nn \ior_map_inline:Nn \ior_get:NN \ior_map_inline:Nn

Applies the \textit{inline function} to each set of \textit{lines} obtained by calling \ior_get:NN until reaching the end of the file. \TeX\ ignores any trailing new-line marker from the file it reads. The \textit{inline function} should consist of code which receives the \textit{line} as \texttt{#1}.

\ior_str_map_inline:Nn \ior_str_map_inline:Nn \ior_str_map_variable:NNn

Applies the \textit{inline function} to every \textit{line} in the \textit{stream}. The material is read from the \textit{stream} as a series of tokens with category code 12 (other), with the exception of space characters which are given category code 10 (space). The \textit{inline function} should consist of code which receives the \textit{line} as \texttt{#1}. Note that \TeX\ removes trailing space and tab characters (character codes 32 and 9) from every line upon input. \TeX\ also ignores any trailing new-line marker from the file it reads.

\ior_map_variable:NNn \ior_map_variable:NNn \ior_map_break:

For each set of \textit{lines} obtained by calling \ior_get:NN until reaching the end of the file, stores the \textit{lines} in the \textit{tl var} then applies the \textit{code}. The \textit{code} will usually make use of the \textit{variable}, but this is not enforced. The assignments to the \textit{variable} are local. Its value after the loop is the last set of \textit{lines}, or its original value if the \textit{stream} is empty. \TeX\ ignores any trailing new-line marker from the file it reads. This function is typically faster than \ior_str_map_inline:Nn.

\ior_str_map_variable:NNn \ior_str_map_variable:NNn \ior_map_break:

For each \textit{line} in the \textit{stream}, stores the \textit{line} in the \textit{variable} then applies the \textit{code}. The material is read from the \textit{stream} as a series of tokens with category code 12 (other), with the exception of space characters which are given category code 10 (space). The \textit{code} will usually make use of the \textit{variable}, but this is not enforced. The assignments to the \textit{variable} are local. Its value after the loop is the last \textit{line}, or its original value if the \textit{stream} is empty. Note that \TeX\ removes trailing space and tab characters (character codes 32 and 9) from every line upon input. \TeX\ also ignores any trailing new-line marker from the file it reads. This function is typically faster than \ior_str_map_inline:Nn.

\ior_map_break:

Used to terminate a \ior_map... function before all lines from the \textit{stream} have been processed. This normally takes place within a conditional statement, for example

\begin{verbatim}
\ior_map_inline:Nn \l_my_ior
{\str_if_eq:nnTF { #1 } { bingo }{ \ior_map_break: }
{ % Do something useful }
}
\end{verbatim}

Use outside of a \ior_map... scenario leads to low level \TeX\ errors.

\TeX\ hackers note: When the mapping is broken, additional tokens may be inserted before further items are taken from the input stream. This depends on the design of the mapping function.
\ior_map_break:n \ior_map_break:n \{\langle code\rangle\}

Used to terminate a \ior_map... function before all lines in the \langle stream \rangle have been processed, inserting the \langle code \rangle after the mapping has ended. This normally takes place within a conditional statement, for example

\ior_map_inline:Nn \l_my_ior
{\str_if_eq:nnTF { #1 } { bingo }\{ \ior_map_break:n \{ <code> \} \}
{\% Do something useful\}
}

Use outside of a \ior_map... scenario leads to low level \TeX{} errors.

\TeX{}hackers note: When the mapping is broken, additional tokens may be inserted before the \langle code \rangle is inserted into the input stream. This depends on the design of the mapping function.

\ior_if_eof_p:N  
\ior_if_eof:NTF

Tests if the end of a file \langle stream \rangle has been reached during a reading operation. The test also returns a true value if the \langle stream \rangle is not open.

1.2 Writing to files

\iow_now:Nn  \iow_now:Nn \{\langle tokens\rangle\}

This functions writes \langle tokens \rangle to the specified \langle stream \rangle immediately \(i.e.\) the write operation is called on expansion of \iow_now:Nn).

\iow_log:n  \iow_log:n \{\langle tokens\rangle\}

This function writes the given \langle tokens \rangle to the log (transcript) file immediately: it is a dedicated version of \iow_now:Nn.

\iow_term:n  \iow_term:n \{\langle tokens\rangle\}

This function writes the given \langle tokens \rangle to the terminal file immediately: it is a dedicated version of \iow_now:Nn.
\iow_shipout\texttt{N} \{\texttt{stream}\} \{\{\texttt{tokens}\}\}

This functions writes \{\texttt{tokens}\} to the specified \{\texttt{stream}\} when the current page is finalised (\textit{i.e.} at shipout). The \texttt{x}-type variants expand the \{\texttt{tokens}\} at the point where the function is used but \textit{not} when the resulting tokens are written to the \{\texttt{stream}\} (\textit{cf.} \texttt{\iow_shipout\_x:Nn}).

\textbf{\TeXhacks note:} When using expl3 with a format other than \LaTeX, new line characters inserted using \texttt{\iow_newline}: or using the line-wrapping code \texttt{\iow\_wrap:nnnN} are not recognized in the argument of \texttt{\iow\_shipout\_Nn}. This may lead to the insertion of additional unwanted line-breaks.

\iow_shipout\_x\texttt{N} \{\texttt{stream}\} \{\{\texttt{tokens}\}\}

This functions writes \{\texttt{tokens}\} to the specified \{\texttt{stream}\} when the current page is finalised (\textit{i.e.} at shipout). The \{\texttt{tokens}\} are expanded at the time of writing in addition to any expansion when the function is used. This makes these functions suitable for including material finalised during the page building process (such as the page number integer).

\textbf{\TeXhacks note:} This is a wrapper around the \TeX\ primitive \texttt{\write}. When using expl3 with a format other than \LaTeX, new line characters inserted using \texttt{\iow_newline}: or using the line-wrapping code \texttt{\iow\_wrap:nnnN} are not recognized in the argument of \texttt{\iow\_shipout\_Nn}. This may lead to the insertion of additional unwanted line-breaks.

\iow_char\texttt{N} \texttt{*}

\texttt{\iow_char:N} \{\texttt{char}\}

Inserts \{\texttt{char}\} into the output stream. Useful when trying to write difficult characters such as \%, \{, \}, \texttt{etc}. in messages, for example:

\begin{verbatim}
\iow_now:Nx \g_my_iow { \iow_char:N \{ text \iow_char:N \} }
\end{verbatim}

The function has no effect if writing is taking place without expansion (\textit{e.g.} in the second argument of \texttt{\iow_now:Nn}).

\iow_newline: \texttt{*}

Function to add a new line within the \{\texttt{tokens}\} written to a file. The function has no effect if writing is taking place without expansion (\textit{e.g.} in the second argument of \texttt{\iow\_now:Nn}).

\textbf{\TeXhacks note:} When using expl3 with a format other than \LaTeX, the character inserted by \texttt{\iow\_newline}: is not recognized by \TeX, which may lead to the insertion of additional unwanted line-breaks. This issue only affects \texttt{\iow\_shipout\_Nn}, \texttt{\iow\_shipout\_x:Nn} and direct uses of primitive operations.
1.3 Wrapping lines in output

\iow_wrap:nnnN \iow_wrap:nnxnN

This function wraps the \langle text \rangle to a fixed number of characters per line. At the start of each line which is wrapped, the \langle run-on text \rangle is inserted. The line character count targeted is the value of \l_iow_line_count_int minus the number of characters in the \langle run-on text \rangle for all lines except the first, for which the target number of characters is simply \l_iow_line_count_int since there is no run-on text. The \langle text \rangle and \langle run-on text \rangle are exhaustively expanded by the function, with the following substitutions:

- \\ or \iow_newline: may be used to force a new line,
- \_ may be used to represent a forced space (for example after a control sequence),
- \#, \%, \{, \}, \- may be used to represent the corresponding character,
- \iow_allow_break: may be used to allow a line-break without inserting a space (this is experimental),
- \iow_indent:n may be used to indent a part of the \langle text \rangle (not the \langle run-on text \rangle).

Additional functions may be added to the wrapping by using the \langle set up \rangle, which is executed before the wrapping takes place: this may include overriding the substitutions listed.

Any expandable material in the \langle text \rangle which is not to be expanded on wrapping should be converted to a string using \token_to_str:N, \tl_to_str:n, \tl_to_str:N, etc.

The result of the wrapping operation is passed as a braced argument to the \langle function \rangle, which is typically a wrapper around a write operation. The output of \iow_wrap:nnnN (i.e. the argument passed to the \langle function \rangle) consists of characters of category “other” (category code 12), with the exception of spaces which have category “space” (category code 10). This means that the output does not expand further when written to a file.

\TeXhacks note: Internally, \iow_wrap:nnnN carries out an x-type expansion on the \langle text \rangle to expand it. This is done in such a way that \exp_not:N or \exp_not:n could be used to prevent expansion of material. However, this is less conceptually clear than conversion to a string, which is therefore the supported method for handling expandable material in the \langle text \rangle.

\iow_indent:n \iow_indent:n \{(text)\}

In the first argument of \iow_wrap:nnnN (for instance in messages), indents \langle text \rangle by four spaces. This function does not cause a line break, and only affects lines which start within the scope of the \langle text \rangle. In case the indented \langle text \rangle should appear on separate lines from the surrounding text, use \\ to force line breaks.

\l_iow_line_count_int

The maximum number of characters in a line to be written by the \iow_wrap:nnnN function. This value depends on the \TeX{} system in use: the standard value is 78, which is typically correct for unmodified \TeX{}live and MiK\TeX{} systems.
1.4 Constant input–output streams, and variables

\g_tmpa_ior
\g_tmpb_ior
Rev: 2017-12-11

Scratch input stream for global use. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\c_log_iow
\c_term_iow

Constant output streams for writing to the log and to the terminal (plus the log), respectively.

\g_tmpa_iow
\g_tmpb_iow
Rev: 2017-12-11

Scratch output stream for global use. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

1.5 Primitive conditionals

\if_eof:w \langle stream \rangle \langle true code \rangle \else: \langle false code \rangle \fi:
Tests if the \langle stream \rangle returns “end of file”, which is true for non-existent files. The \else: branch is optional.

\TeXhackers note: This is the \TeX primitive \texttt{\textbackslash ifeof}.

2 File operation functions

\g_file_curr_dir_str
\g_file_curr_name_str
\g_file_curr_ext_str
Rev: 2017-06-21

Contain the directory, name and extension of the current file. The directory is empty if the file was loaded without an explicit path (\textit{i.e.} if it is in the \TeX search path), and does not end in / other than the case that it is exactly equal to the root directory. The \langle name \rangle and \langle ext \rangle parts together make up the file name, thus the \langle name \rangle part may be thought of as the “job name” for the current file. Note that \TeX does not provide information on the \langle ext \rangle part for the main (top level) file and that this file always has an empty \langle dir \rangle component. Also, the \langle name \rangle here will be equal to \texttt{\textbackslash c\_sys\_jobname\_str}, which may be different from the real file name (if set using \texttt{--jobname}, for example).
Each entry is the path to a directory which should be searched when seeking a file. Each path can be relative or absolute, and should not include the trailing slash. The entries are not expanded when used so may contain active characters but should not feature any variable content. Spaces need not be quoted.

\TeXhackers note: When working as a package in \LaTeX 2ε, expl3 will automatically append the current \input@path to the set of values from \l_file_search_path_seq.

\file_if_exist:nTF {{\textit{file name}}} {{\textit{true code}}} {{\textit{false code}}}

Searches for \textit{file name} using the current \TeX search path and the additional paths controlled by \l_file_search_path_seq.

\file_get:nnN {\textit{filename}} {\textit{setup}} {\textit{tl}}
\file_get:nnNTF {\textit{filename}} {\textit{setup}} {\textit{tl}} {{\textit{true code}}} {{\textit{false code}}}

Defines \textit{tl} to the contents of \textit{filename}. Category codes may need to be set appropriately via the \textit{setup} argument. The non-branching version sets the \textit{tl} to \q_no_value if the file is not found. The branching version runs the \textit{true code} after the assignment to \textit{tl} if the file is found, and \textit{false code} otherwise.

\file_get_full_name:nN
\file_get_full_name:VN
\file_get_full_name:nNTF
\file_get_full_name:VNTF

Searches for \textit{file name} in the path as detailed for \file_if_exist:nTF, and if found sets the \textit{tl var} the fully-qualified name of the file, \ie the path and file name. This includes an extension .\textit{tex} when the given \textit{file name} has no extension but the file found has that extension. In the non-branching version, the \textit{tl var} will be set to \q_no_value in the case that the file does not exist.

\file_full_name:n
\file_full_name:V

Searches for \textit{file name} in the path as detailed for \file_if_exist:nTF, and if found leaves the fully-qualified name of the file, \ie the path and file name, in the input stream. This includes an extension .\textit{tex} when the given \textit{file name} has no extension but the file found has that extension. If the file is not found on the path, the expansion is empty.
Parses the \textit{full name} and splits it into three parts, each of which is returned by setting the appropriate local string variable:

- The \textit{dir}: everything up to the last / (path separator) in the \textit{file path}. As with system PATH variables and related functions, the \textit{dir} does not include the trailing / unless it points to the root directory. If there is no path (only a file name), \textit{dir} is empty.

- The \textit{name}: everything after the last / up to the last ., where both of those characters are optional. The \textit{name} may contain multiple . characters. It is empty if \textit{full name} consists only of a directory name.

- The \textit{ext}: everything after the last . (including the dot). The \textit{ext} is empty if there is no . after the last /.

This function does not expand the \textit{full name} before turning it to a string. It assume that the \textit{full name} either contains no quote (") characters or is surrounded by a pair of quotes.

Searches for \textit{file name} using the current TeX search path and the additional paths controlled by \_\_file_search_path_seq. It then expands to leave the hexadecimal dump of the file content in the input stream. The file is read as bytes, which means that in contrast to most TeX behaviour there will be a difference in result depending on the line endings used in text files. The same file will produce the same result between different engines: the algorithm used is the same in all cases. When the file is not found, the result of expansion is empty. The \{\textit{start index}\} and \{\textit{end index}\} values work as described for \texttt{str_range:nnn}.

Sets the \textit{tl var} to the result of applying \texttt{file_hex_dump:n//file_hex_dump:nnn} to the \textit{file}. If the file is not found, the \{\textit{tl var}\} will be set to \texttt{q_no_value}.

Searches for \textit{file name} using the current TeX search path and the additional paths controlled by \_\_file_search_path_seq. It then expands to leave the MD5 sum generated from the contents of the file in the input stream. The file is read as bytes, which means that in contrast to most TeX behaviour there will be a difference in result depending on the line endings used in text files. The same file will produce the same result between different engines: the algorithm used is the same in all cases. When the file is not found, the result of expansion is empty.

Sets the \{\textit{tl var}\} to the result of applying \texttt{file_mdfive_hash:n//file_mdfive_hash:nnn} to the \textit{file}. If the file is not found, the \{\textit{tl var}\} will be set to \texttt{q_no_value}.
\file_size:n \{(file name)\}

Searches for (file name) using the current \TeX{} search path and the additional paths controlled by \_file_search_path_seq. It then expands to leave the size of the file in bytes in the input stream. When the file is not found, the result of expansion is empty.

\file_size:nN \{(file name)\} (tl var)

Sets the (tl var) to the result of applying \file_size:n to the (file). If the file is not found, the (tl var) will be set to \q_no_value. This is not available in older versions of X\TeX{}.

\file_timestamp:n \{(file name)\}

Searches for (file name) using the current \TeX{} search path and the additional paths controlled by \_file_search_path_seq. It then expands to leave the modification timestamp of the file in the input stream. The timestamp is of the form D:\{year\}\{month\}\{day\}\{hour\}\{minute\}\{second\}\{offset\}, where the latter may be Z (UTC) or (plus-minus)\{hours\}\{minutes\}. When the file is not found, the result of expansion is empty. This is not available in older versions of X\TeX{}.

\file_timestamp:nN \{(file name)\} (tl var)

Sets the (tl var) to the result of applying \file_timestamp:n to the (file). If the file is not found, the (tl var) will be set to \q_no_value. This is not available in older versions of X\TeX{}.

\file_compare_timestamp:nNn \{(file-1)\} (comparator) \{(file-2)\} \{(true code)\} \{(false code)\}

Compares the file stamps on the two (files) as indicated by the (comparator), and inserts either the (true code) or (false case) as required. A file which is not found is treated as older than any file which is found. This allows for example the construct

\file_compare_timestamp:nNnT \{ source-file \} > \{ derived-file \}

{ % Code to regenerate derived file
}

to work when the derived file is entirely absent. The timestamp of two absent files is regarded as different. This is not available in older versions of X\TeX{}.

\file_input:n \{(file name)\}

Searches for (file name) in the path as detailed for \file_if_exist:nTF, and if found reads in the file as additional \LaTeX{} source. All files read are recorded for information and the file name stack is updated by this function. An error is raised if the file is not found.
\file_if_exist_input:n \file_if_exist_input:n \file_if_exist_input:n {⟨file name⟩} \file_if_exist_input:nF \file_if_exist_input:nF {⟨false code⟩}

Searches for ⟨file name⟩ using the current \TeX\ search path and the additional paths controlled by \file_path_include:n. If found then reads in the file as additional \LaTeX\ source as described for \file_input:n, otherwise inserts the ⟨false code⟩. Note that these functions do not raise an error if the file is not found, in contrast to \file_input:n.

\file_input_stop:

\file_input_stop:

Ends the reading of a file started by \file_input:n or similar before the end of the file is reached. Where the file reading is being terminated due to an error, \msg_-critical:nn(nn) should be preferred.

\TeX\ hackers note: This function must be used on a line on its own: \TeX\ reads files line-by-line and so any additional tokens in the “current” line will still be read.

This is also true if the function is hidden inside another function (which will be the normal case), i.e., all tokens on the same line in the source file are still processed. Putting it on a line by itself in the definition doesn’t help as it is the line where it is used that counts!

\file_show_list:
\file_log_list:

These functions list all files loaded by \LATEX\ commands that populate \@filelist or by \file_input:n. While \file_show_list: displays the list in the terminal, \file_log_list: outputs it to the log file only.
Part XX
The l3skip package
Dimensions and skips

I3TPeX3 provides two general length variables: \texttt{dim} and \texttt{skip}. Lengths stored as \texttt{dim} variables have a fixed length, whereas \texttt{skip} lengths have a rubber (stretch/shrink) component. In addition, the \texttt{muskip} type is available for use in math mode: this is a special form of \texttt{skip} where the lengths involved are determined by the current math font (in \texttt{mu}). There are common features in the creation and setting of length variables, but for clarity the functions are grouped by variable type.

1 Creating and initialising \texttt{dim} variables

\begin{verbatim}
\dim_new:N \dim_new:cn
\dim_const:Nn \dim_const:cn \{\texttt{dimension expression}\}
\dim_zero:N \dim_zero:cn \dim_gzero:N \dim_gzero:cn
\dim_zero_new:N \dim_zero_new:cn \dim_gzero_new:N \dim_gzero_new:cn
\dim_if_exist_p:N \dim_if_exist_p:cn \dim_if_exist:NTF \dim_if_exist:cNTF
\end{verbatim}

\begin{itemize}
\item \texttt{\dim_new:N \dim_new:cn} \hspace{4cm} \texttt{\dim_new:N \{\texttt{dimension}\}}
\end{itemize}

Creates a new \texttt{\{\texttt{dimension}\}} or raises an error if the name is already taken. The declaration is global. The \texttt{\{\texttt{dimension}\}} is initially equal to 0 pt.

\begin{itemize}
\item \texttt{\dim_const:Nn \dim_const:cn \{\texttt{dimension expression}\}}
\end{itemize}

Creates a new constant \texttt{\{\texttt{dimension}\}} or raises an error if the name is already taken. The value of the \texttt{\{\texttt{dimension}\}} is set globally to the \texttt{\{\texttt{dimension expression}\}}.

\begin{itemize}
\item \texttt{\dim_zero:N \dim_zero:cn \dim_gzero:N \dim_gzero:cn}
\end{itemize}

Sets \texttt{\{\texttt{dimension}\}} to 0 pt.

\begin{itemize}
\item \texttt{\dim_zero_new:N \dim_zero_new:cn \dim_gzero_new:N \dim_gzero_new:cn}
\end{itemize}

Ensures that the \texttt{\{\texttt{dimension}\}} exists globally by applying \texttt{\dim_new:N} if necessary, then applies \texttt{\dim\_\{\texttt{g}\}zero:N} to leave the \texttt{\{\texttt{dimension}\}} set to zero.

\begin{itemize}
\item \texttt{\dim_if_exist_p:N \dim_if_exist_p:cn \dim_if_exist:NTF \dim_if_exist:cNTF}
\end{itemize}

Tests whether the \texttt{\{\texttt{dimension}\}} is currently defined. This does not check that the \texttt{\{\texttt{dimension}\}} really is a dimension variable.
2 Setting dim variables

\dim_add:Nn \dim_add:cn \dim_gadd:Nn \dim_gadd:cn
Adds the result of the \langle dimension expression \rangle to the current content of the \langle dimension \rangle.
Updated: 2011-10-22

\dim_set:Nn \dim_set:cn \dim_gset:Nn \dim_gset:cn
Sets \langle dimension \rangle to the value of \langle dimension expression \rangle, which must evaluate to a length with units.
Updated: 2011-10-22

\dim_set_eq:NN \dim_set_eq:cn \dim_gset_eq:NN
Sets the content of \langle dimension_1 \rangle equal to that of \langle dimension_2 \rangle.

\dim_sub:Nn \dim_sub:cn \dim_gsub:Nn \dim_gsub:cn
Subtracts the result of the \langle dimension expression \rangle from the current content of the \langle dimension \rangle.
Updated: 2011-10-22

3 Utilities for dimension calculations

\dim_abs:n \dim_abs:n* \dim_abs:n \\dim_abs:n \langle dimexpr \rangle
Converts the \langle dimexpr \rangle to its absolute value, leaving the result in the input stream as a \langle dimension denotation \rangle.
Updated: 2012-09-26

\dim_max:nn \dim_max:nn* \dim_max:nn \\dim_max:nn \langle dimexpr_1 \rangle \langle dimexpr_2 \rangle
\dim_min:nn \dim_min:nn* \dim_min:nn \\dim_min:nn \langle dimexpr_1 \rangle \langle dimexpr_2 \rangle
Evaluates the two \langle dimension expressions \rangle and leaves either the maximum or minimum value in the input stream as appropriate, as a \langle dimension denotation \rangle.
Updated: 2012-09-26
\texttt{dim\_ratio:nn} \texttt{(dimexpr\(_1\)) \texttt{(dimexpr\(_2\))}}

Parses the two \textit{dimension expressions} and converts the ratio of the two to a form suitable for use inside a \textit{dimension expression}. This ratio is then left in the input stream, allowing syntax such as

\begin{verbatim}
\dim_set:Nn \l_my_dim
{ 10 pt * \dim\_ratio:nn { 5 pt } { 10 pt } }
\end{verbatim}

The output of \texttt{dim\_ratio:nn} on full expansion is a ratio expression between two integers, with all distances converted to scaled points. Thus

\begin{verbatim}
\tl_set:Nx \l_my_tl { \dim\_ratio:nn { 5 pt } { 10 pt } }
\tl_show:N \l_my_tl
\end{verbatim}

displays \texttt{327680/655360} on the terminal.

4 Dimension expression conditionals

\texttt{dim\_compare:p:nNn} \texttt{\texttt{dim\_compare:nNnTF}}

\texttt{dim\_compare:p:nNn} \texttt{(dimexpr\(_1\)) \texttt{(relation)} \texttt{(dimexpr\(_2\))}}
\texttt{dim\_compare:nNnTF} \texttt{\{(dimexpr\(_1\)) \texttt{(relation)} \texttt{(dimexpr\(_2\))}\}}
\texttt{\{(true code\)} \texttt{\{(false code\)}}

This function first evaluates each of the \textit{dimension expressions} as described for \texttt{dim\_eval:n}. The two results are then compared using the \textit{relation}:

\begin{center}
\begin{tabular}{ccc}
Equal & = \\
Greater than & \textgreater \\
Less than & \textless \\
\end{tabular}
\end{center}

This function is less flexible than \texttt{dim\_compare:nTF} but around 5 times faster.
This function evaluates the \textit{dimension expressions} as described for \texttt{\dim_eval:n} and compares consecutive result using the corresponding \textit{relations}, namely it compares \texttt{\dimexpr_1} and \texttt{\dimexpr_2} using the \texttt{\relation_1}, then \texttt{\dimexpr_2} and \texttt{\dimexpr_3} using the \texttt{\relation_2}, until finally comparing \texttt{\dimexpr_N} and \texttt{\dimexpr_{N+1}} using the \texttt{\relation_N}. The test yields \texttt{true} if all comparisons are \texttt{true}. Each \textit{dimension expression} is evaluated only once, and the evaluation is lazy, in the sense that if one comparison is \texttt{false}, then no other \textit{dimension expression} is evaluated and no other comparison is performed. The \textit{relations} can be any of the following:

\begin{align*}
\text{Equal} &\quad = \quad \text{or} \quad == \\
\text{Greater than or equal to} &\quad \geq \\
\text{Greater than} &\quad > \\
\text{Less than or equal to} &\quad \leq \\
\text{Less than} &\quad < \\
\text{Not equal} &\quad \neq
\end{align*}

This function is more flexible than \texttt{\dim_compare:nNnTF} but around 5 times slower.
\dim_case:nn \dim_case:nnTF \{(test dimension expression)\} \
\dim_case:nnTF \{ 
\{\dimexpr case_1\} \{\code case_1\} 
\{\dimexpr case_2\} \{\code case_2\} 
\ldots 
\{\dimexpr case_n\} \{\code case_n\} 
\} 
\{(true code)\} 
\{(false code)\}

This function evaluates the \textit{(test dimension expression)} and compares this in turn to each of the \textit{(dimension expression cases)}. If the two are equal then the associated \textit{(code)} is left in the input stream and other cases are discarded. If any of the cases are matched, the \textit{(true code)} is also inserted into the input stream (after the code for the appropriate case), while if none match then the \textit{(false code)} is inserted. The function \texttt{\dim_case:nn}, which does nothing if there is no match, is also available. For example

\begin{verbatim}
\dim_set:Nn \l_tmpa_dim { 5 pt } 
\dim_case:nnF 
\{ 2 \l_tmpa_dim \} 
\{ 
\{ 5 pt \} \{ Small \} 
\{ 4 pt + 6 pt \} \{ Medium \} 
\{ - 10 pt \} \{ Negative \} 
\} 
\{ No idea! \}
\end{verbatim}

leaves “Medium” in the input stream.

5 Dimension expression loops

\dim_do_until:nNnn \dim_do_until:nNnn \{(dimexpr_1)\} \{relation\} \{(dimexpr_2)\} \{(code)\}

Places the \textit{(code)} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \textit{(dimension expressions)} as described for \texttt{\dim_compare:nNnTF}. If the test is \textit{false} then the \textit{(code)} is inserted into the input stream again and a loop occurs until the \textit{(relation)} is \textit{true}.

\dim_do_while:nNnn \dim_do_while:nNnn \{(dimexpr_1)\} \{relation\} \{(dimexpr_2)\} \{(code)\}

Places the \textit{(code)} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \textit{(dimension expressions)} as described for \texttt{\dim_compare:nNnTF}. If the test is \textit{true} then the \textit{(code)} is inserted into the input stream again and a loop occurs until the \textit{(relation)} is \textit{false}.

\dim_until_do:nNnn \dim_until_do:nNnn \{(dimexpr_1)\} \{relation\} \{(dimexpr_2)\} \{(code)\}

Evaluates the relationship between the two \textit{(dimension expressions)} as described for \texttt{\dim_compare:nNnTF}, and then places the \textit{(code)} in the input stream if the \textit{(relation)} is \textit{false}. After the \textit{(code)} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \textit{true}.
\dim_step_function:nnn \dim_step_function:nnn \dim_step_function:nnn \dim_step_function:nnn \dim_step_function:nnn \dim_step_function:nnn

\dim_step_function:nnn \{\textsl{initial value}\} \{\textsl{step}\} \{\textsl{final value}\} \{\textsl{function}\}

This function first evaluates the \textsl{initial value}, \textsl{step} and \textsl{final value}, all of which should be dimension expressions. The \textsl{function} is then placed in front of each \textsl{value} from the \textsl{initial value} to the \textsl{final value} in turn (using \textsl{step} between each \textsl{value}). The \textsl{step} must be non-zero. If the \textsl{step} is positive, the loop stops when the \textsl{value} becomes larger than the \textsl{final value}. If the \textsl{step} is negative, the loop stops when the \textsl{value} becomes smaller than the \textsl{final value}. The \textsl{function} should absorb one argument.

\dim_step_inline:nnn \\dim_step_inline:nnn \dim_step_inline:nnn \dim_step_inline:nnn \dim_step_inline:nnn

\dim_step_inline:nnn \{\textsl{initial value}\} \{\textsl{step}\} \{\textsl{final value}\} \{\textsl{code}\}

This function first evaluates the \textsl{initial value}, \textsl{step} and \textsl{final value}, all of which should be dimension expressions. Then for each \textsl{value} from the \textsl{initial value} to the \textsl{final value} in turn (using \textsl{step} between each \textsl{value}), the \textsl{code} is inserted into the input stream with \#1 replaced by the current \textsl{value}. Thus the \textsl{code} should define a function of one argument (\#1).
This function first evaluates the \( \langle \text{initial value} \rangle \), \( \langle \text{step} \rangle \) and \( \langle \text{final value} \rangle \), all of which should be dimension expressions. Then for each \( \langle \text{value} \rangle \) from the \( \langle \text{initial value} \rangle \) to the \( \langle \text{final value} \rangle \) in turn (using \( \langle \text{step} \rangle \) between each \( \langle \text{value} \rangle \)), the \( \langle \text{code} \rangle \) is inserted into the input stream, with the \( \langle \text{tl var} \rangle \) defined as the current \( \langle \text{value} \rangle \). Thus the \( \langle \text{code} \rangle \) should make use of the \( \langle \text{tl var} \rangle \).

### 7 Using \texttt{dim} expressions and variables

\texttt{\dim_eval:n} \( \{ \langle \text{dimension expression} \rangle \} \)

Evaluates the \( \langle \text{dimension expression} \rangle \), expanding any dimensions and token list variables within the \( \langle \text{expression} \rangle \) to their content (without requiring \texttt{\dim_use:N} or \texttt{\tl_use:N}) and applying the standard mathematical rules. The result of the calculation is left in the input stream as a \( \langle \text{dimension denotation} \rangle \) after two expansions. This is expressed in points (pt), and requires suitable termination if used in a \TeX{}-style assignment as it is not an \( \langle \text{internal dimension} \rangle \).

\texttt{\dim_sign:n} \( \{ \langle \text{dimexpr} \rangle \} \)

Evaluates the \( \langle \text{dimexpr} \rangle \) then leaves 1 or 0 or \(-1\) in the input stream according to the sign of the result.

\texttt{\dim_use:N} \( \langle \text{dimension} \rangle \) \( \texttt{\dim_use:c} \)

Recover the content of a \( \langle \text{dimension} \rangle \) and places it directly in the input stream. An error is raised if the variable does not exist or if it is invalid. Can be omitted in places where a \( \langle \text{dimension} \rangle \) is required (such as in the argument of \texttt{\dim_eval:n}).

\TeX{}hackers note: \texttt{\dim_use:N} is the \TeX{} primitive \texttt{\the}: this is one of several \LaTeX{}3 names for this primitive.

\texttt{\dim_to_decimal:n} \( \{ \langle \text{dimexpr} \rangle \} \)

Evaluates the \( \langle \text{dimension expression} \rangle \), and leaves the result, expressed in points (pt) in the input stream, with \textit{no units}. The result is rounded by \TeX{} to four or five decimal places. If the decimal part of the result is zero, it is omitted, together with the decimal marker.

For example

\texttt{\dim_to_decimal:n \{ 1bp \}}

leaves 1.00374 in the input stream, \textit{i.e.} the magnitude of one “big point” when converted to (\TeX{}) points.
\dim_to_decimal_in_bp:n \{\langle\text{dimexpr}\rangle\}

Evaluates the \langle\text{dimension expression}\rangle, and leaves the result, expressed in big points (bp) in the input stream, with \textit{no units}. The result is rounded by \TeX to four or five decimal places. If the decimal part of the result is zero, it is omitted, together with the decimal marker.

For example

\dim_to_decimal_in_bp:n \{1\text{pt}\}

leaves 0.99628 in the input stream, \textit{i.e.} the magnitude of one (\TeX) point when converted to big points.

\dim_to_decimal_in_sp:n \{\langle\text{dimexpr}\rangle\}

Evaluates the \langle\text{dimension expression}\rangle, and leaves the result, expressed in scaled points (sp) in the input stream, with \textit{no units}. The result is necessarily an integer.

\dim_to_decimal_in_unit:nn \{\langle\text{dimexpr}_1\rangle\} \{\langle\text{dimexpr}_2\rangle\}

Evaluates the \langle\text{dimension expressions}\rangle, and leaves the value of \langle\text{dimexpr}_1\rangle, expressed in a unit given by \langle\text{dimexpr}_2\rangle, in the input stream. The result is a decimal number, rounded by \TeX to four or five decimal places. If the decimal part of the result is zero, it is omitted, together with the decimal marker.

For example

\dim_to_decimal_in_unit:nn \{1\text{bp}\} \{1\text{mm}\}

leaves 0.35277 in the input stream, \textit{i.e.} the magnitude of one big point when converted to millimetres.

Note that this function is not optimised for any particular output and as such may give different results to \dim_to_decimal_in_bp:n or \dim_to_decimal_in_sp:n. In particular, the latter is able to take a wider range of input values as it is not limited by the ability to calculate a ratio using \varepsilon-\TeX primitives, which is required internally by \dim_to_decimal_in_unit:nn.

\dim_to_fp:n \{\langle\text{dimexpr}\rangle\}

Expands to an internal floating point number equal to the value of the \langle\text{dimexpr}\rangle in pt. Since dimension expressions are evaluated much faster than their floating point equivalent, \dim_to_fp:n can be used to speed up parts of a computation where a low precision and a smaller range are acceptable.

8 Viewing dim variables

\dim_show:N \langle\text{dimension}\rangle

\dim_show:c

Displays the value of the \langle\text{dimension}\rangle on the terminal.
\dim_show:n \{\langle\text{dimension expression}\rangle\}

Displays the result of evaluating the \langle\text{dimension expression}\rangle on the terminal.

\dim_log:N \dim_log:c

\dim_log:n \{\langle\text{dimension expression}\rangle\}

\dim_log:N \dim_log:c

\dim_log:n \{\langle\text{dimension expression}\rangle\}

9 Constant dimensions

\c_max_dim

The maximum value that can be stored as a dimension. This can also be used as a component of a skip.

\c_zero_dim

A zero length as a dimension. This can also be used as a component of a skip.

10 Scratch dimensions

\l_tmpa_dim \l_tmpb_dim

Scratch dimension for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_dim \g_tmpb_dim

Scratch dimension for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

11 Creating and initialising \texttt{skip} variables

\skip_new:N \skip_new:c

\skip_new:N \{\langle\text{skip}\rangle\}

Creates a new \langle\text{skip}\rangle or raises an error if the name is already taken. The declaration is global. The \langle\text{skip}\rangle is initially equal to 0pt.
Creates a new constant \( \langle \text{skip} \rangle \) or raises an error if the name is already taken. The value of the \( \langle \text{skip} \rangle \) is set globally to the \( \langle \text{skip expression} \rangle \).

Sets \( \langle \text{skip} \rangle \) to 0 pt.

Ensures that the \( \langle \text{skip} \rangle \) exists globally by applying \texttt{\textbackslash skip\_new:N} if necessary, then applies \texttt{\textbackslash skip\_\textit{(g)zero:N}} to leave the \( \langle \text{skip} \rangle \) set to zero.

Tests whether the \( \langle \text{skip} \rangle \) is currently defined. This does not check that the \( \langle \text{skip} \rangle \) really is a skip variable.

12 Setting skip variables

Adds the result of the \( \langle \text{skip expression} \rangle \) to the current content of the \( \langle \text{skip} \rangle \).

Sets \( \langle \text{skip} \rangle \) to the value of \( \langle \text{skip expression} \rangle \), which must evaluate to a length with units and may include a rubber component (for example 1 cm plus 0.5 cm).

Sets the content of \( \langle \text{skip}_1 \rangle \) equal to that of \( \langle \text{skip}_2 \rangle \).

Subtracts the result of the \( \langle \text{skip expression} \rangle \) from the current content of the \( \langle \text{skip} \rangle \).
13 Skip expression conditionals

\skip_if_eq_p:nn \skip_if_eq_p:nn \{\langle skipexpr_1\rangle\} \{\langle skipexpr_2\rangle\}
\skip_if_eq:nTF \{\langle skipexpr_1\rangle\} \{\langle skipexpr_2\rangle\}
\{\langle true code\rangle\} \{\langle false code\rangle\}

This function first evaluates each of the \langle skip expressions\rangle as described for \skip_eval:n. The two results are then compared for exact equality, i.e. both the fixed and rubber components must be the same for the test to be true.

\skip_if_finite_p:n \skip_if_finite_p:n \{\langle skipexpr\rangle\}
\skip_if_finite:nTF \{\langle skipexpr\rangle\} \{\langle true code\rangle\} \{\langle false code\rangle\}

Evaluates the \langle skip expression\rangle as described for \skip_eval:n, and then tests if all of its components are finite.

14 Using skip expressions and variables

\skip_eval:n \{\langle skip expression\rangle\}

Evaluates the \langle skip expression\rangle, expanding any skips and token list variables within the \langle expression\rangle to their content (without requiring \skip_use:N/\tl_use:N) and applying the standard mathematical rules. The result of the calculation is left in the input stream as a \langle glue denotation\rangle after two expansions. This is expressed in points (pt), and requires suitable termination if used in a \TeX-style assignment as it is not an \langle internal glue\rangle.

\skip_use:N \{\langle skip\rangle\}

Recovers the content of a \langle skip\rangle and places it directly in the input stream. An error is raised if the variable does not exist or if it is invalid. Can be omitted in places where a \langle dimension\rangle or \langle skip\rangle is required (such as in the argument of \skip_eval:n).

\TeXhacksnote: \skip_use:N is the \TeX primitive \the: this is one of several \TeX names for this primitive.

15 Viewing skip variables

\skip_show:N \{\langle skip\rangle\}

Displays the value of the \langle skip\rangle on the terminal.

\skip_show:n \{\langle skip expression\rangle\}

Displays the result of evaluating the \langle skip expression\rangle on the terminal.
\skip_log:N
\skip_log:c

New: 2014-08-22
Updated: 2015-08-03

\skip_log:n

New: 2014-08-22
Updated: 2015-08-07

\skip_log:N \langle \text{skip} \rangle

Writes the value of the \langle \text{skip} \rangle in the log file.

\skip_log:n \{ \langle \text{skip expression} \rangle \}

Writes the result of evaluating the \langle \text{skip expression} \rangle in the log file.

16 Constant skips

\c_max_skip

Updated: 2012-11-02

The maximum value that can be stored as a skip (equal to \c_max_dim in length), with no stretch nor shrink component.

\c_zero_skip

Updated: 2012-11-01

A zero length as a skip, with no stretch nor shrink component.

17 Scratch skips

\l_tmpa_skip
\l_tmpb_skip

Scratch skip for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_skip
\g_tmpb_skip

Scratch skip for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

18 Inserting skips into the output

\skip_horizontal:N
\skip_horizontal:c
\skip_horizontal:n

Updated: 2011-10-22

\skip_horizontal:N \langle \text{skip} \rangle
\skip_horizontal:n \{ \langle \text{skipexpr} \rangle \}

Inserts a horizontal \langle \text{skip} \rangle into the current list. The argument can also be a \langle \text{dim} \rangle.

\TeXhackers note: \skip_horizontal:N is the \TeX primitive \hskip renamed.
\skip_vertical:N \skip_vertical:c \skip_vertical:n

Updated: 2011-10-22

\skip_vertical:N \skip
\skip_vertical:n \{\skipeexpr\}

Inserts a vertical \skip into the current list. The argument can also be a \dim.

\TeXhackers\ note: \skip_vertical:N is the \TeX\ primitive \vskip renamed.

19 Creating and initialising muskip variables

\muskip_new:N \muskip_new:c

Creates a new \muskip or raises an error if the name is already taken. The declaration is global. The \muskip is initially equal to 0mu.

\muskip_const:Nn \muskip_const:cn

New: 2012-03-05

\skipe_zero:N \skipe
\skipe_zero:c \skipe_gzero:N \skipe_gzero:c

\muskip_zero_new:N \muskip_zero_new:c \muskip_gzero_new:N \muskip_gzero_new:c

New: 2012-01-07

\skipe_if_exist_p:N \skipe_if_exist:N \skipe_if_exist:c
\skipe_if_exist:NTF \skipe_if_exist:cN\skipe_if_exist:cT

New: 2012-03-03

20 Setting muskip variables

\muskip_add:Nn \muskip_add:cn \muskip_gadd:Nn \muskip_gadd:cn

Updated: 2011-10-22

\muskip_add:Nn \muskip \{\muskip expression\}

Adds the result of the \muskip expression to the current content of the \muskip.
\texttt{\textbackslash muskip\_set:Nn} \texttt{\textbackslash muskip\_set:cn} \texttt{\textbackslash muskip\_gset:Nn} \texttt{\textbackslash muskip\_gset:cn}
\texttt{\textbackslash muskip\_set_eq:NN} \texttt{\textbackslash muskip\_set_eq:cn} \texttt{\textbackslash muskip\_gset_eq:NN} \texttt{\textbackslash muskip\_gset_eq:cn}
\texttt{\textbackslash muskip\_sub:Nn} \texttt{\textbackslash muskip\_sub:cn} \texttt{\textbackslash muskip\_gsub:Nn} \texttt{\textbackslash muskip\_gsub:cn}
\texttt{\textbackslash muskip\_eval:n} \texttt{\textbackslash muskip\_use:N} \texttt{\textbackslash muskip\_use:c}
\texttt{\textbackslash muskip\_show:N} \texttt{\textbackslash muskip\_show:c}

\textbf{21 Using muskip expressions and variables}

\texttt{\textbackslash muskip\_eval:n} \texttt{\textbackslash muskip\_use:N} \texttt{\textbackslash muskip\_use:c}

\textbf{22 Viewing muskip variables}
\muskip\show:n \muskip\show:n \langle \muskip\ expression \rangle \\
Displays the result of evaluating the \langle \muskip\ expression \rangle on the terminal.

\muskip\log:N \muskip\log:c \\
\muskip\log:N \langle \muskip \rangle \\
Writes the value of the \langle \muskip \rangle in the log file.

\muskip\log:n \muskip\log:n \\
\muskip\log:n \langle \muskip\ expression \rangle \\
Writes the result of evaluating the \langle \muskip\ expression \rangle in the log file.

23 Constant muskips

\c_max_muskip \\
The maximum value that can be stored as a muskip, with no stretch nor shrink component.

\c_zero_muskip \\
A zero length as a muskip, with no stretch nor shrink component.

24 Scratch muskips

\l_tmpa_muskip \l_tmpb_muskip \\
Scratch muskip for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_muskip \g_tmpb_muskip \\
Scratch muskip for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

25 Primitive conditional

\if_dim:w \star \if_dim:w \langle \dimen1 \rangle \langle \relation \rangle \langle \dimen2 \rangle \\
\langle \true\ code \rangle \\
\else: \\
\langle \false \rangle \\
\fi: \\
Compare two dimensions. The \langle \relation \rangle is one of <, = or > with category code 12.

\TeXhackers\ note: This is the \TeX\ primitive \ifdim.
Part XXI

The l3keys package

Key–value interfaces

The key–value method is a popular system for creating large numbers of settings for controlling function or package behaviour. The system normally results in input of the form

\MyModuleSetup{
    key-one = value one,
    key-two = value two
}

or

\MyModuleMacro[
    key-one = value one,
    key-two = value two
]{argument}

for the user.

The high level functions here are intended as a method to create key–value controls. Keys are themselves created using a key–value interface, minimising the number of functions and arguments required. Each key is created by setting one or more properties of the key:

\keys_define:nn { mymodule }
{
    key-one .code:n = code including parameter #1,
    key-two .tl_set:N = \l_mymodule_store_tl
}

These values can then be set as with other key–value approaches:

\keys_set:nn { mymodule }
{
    key-one = value one,
    key-two = value two
}

At a document level, \keys_set:nn is used within a document function, for example

\DeclareDocumentCommand \MyModuleSetup { m }
{ \keys_set:nn { mymodule } { #1 } }
\DeclareDocumentCommand \MyModuleMacro { o m }
{ \group_begin: \keys_set:nn { mymodule } { #1 } % Main code for \MyModuleMacro \group_end: }

183
Key names may contain any tokens, as they are handled internally using \_to-_str:n. As discussed in section 2, it is suggested that the character / is reserved for sub-division of keys into logical groups. Functions and variables are not expanded when creating key names, and so

\_set:Nn \_mymodule_tmp_tl { key }
\keys_define:nn { mymodule }
{
    \_mymodule_tmp_tl .code:n = code
}

creates a key called \_mymodule_tmp_tl, and not one called key.

1 Creating keys

\keys_define:nn \keys_define:nn {⟨module⟩} {⟨keyval list⟩}

Parses the ⟨keyval list⟩ and defines the keys listed there for ⟨module⟩. The ⟨module⟩ name is treated as a string. In practice the ⟨module⟩ should be chosen to be unique to the module in question (unless deliberately adding keys to an existing module).

The ⟨keyval list⟩ should consist of one or more key names along with an associated key property. The properties of a key determine how it acts. The individual properties are described in the following text; a typical use of \keys_define:nn might read

\keys_define:nn { mymodule }
{
    keyname .code:n = Some-code-using-#1,
    keyname .value_required:n = true
}

where the properties of the key begin from the . after the key name.

The various properties available take either no arguments at all, or require one or more arguments. This is indicated in the name of the property using an argument specification. In the following discussion, each property is illustrated attached to an arbitrary ⟨key⟩, which when used may be supplied with a ⟨value⟩. All key definitions are local.

Key properties are applied in the reading order and so the ordering is significant. Key properties which define “actions”, such as .code:n, .tl_set:N, etc., override one another. Some other properties are mutually exclusive, notably .value_required:n and .value_forbidden:n, and so they replace one another. However, properties covering non-exclusive behaviours may be given in any order. Thus for example the following definitions are equivalent.

\keys_define:nn { mymodule }
{
    keyname .code:n = Some-code-using-#1,
    keyname .value_required:n = true
}
\keys_define:nn { mymodule }
{

Note that with the exception of the special `.undefine:` property, all key properties define the key within the current \TeX{} scope.

\begin{verbatim}
keyname .value_required:n = true,
keyname .code:n = Some-code-using-#1
}
\end{verbatim}

\begin{verbatim}
\bool_set:N = \boolean
\bool_set:c
\bool_gset:N
\bool_gset:c
\end{verbatim}

Updated: 2013-07-08

\begin{verbatim}
\bool_set_inverse:N = \boolean
\bool_set_inverse:c
\bool_gset_inverse:N
\bool_gset_inverse:c
\end{verbatim}

Updated: 2013-07-08

\begin{verbatim}
\choice:
\choices:nn = { \choices } \{ \code \}
\end{verbatim}

Updated: 2013-07-10

\begin{verbatim}
\clist_set:N = \comma list variable
\clist_set:c
\clist_gset:N
\clist_gset:c
\end{verbatim}

Updated: 2013-07-10

\begin{verbatim}
\code:n = \{ \code \}
\end{verbatim}

Updated: 2013-07-10

\begin{verbatim}
\cs_set:Np
\cs_set:cp
\cs_set_protected:Np
\cs_set_protected:cp
\end{verbatim}

Updated: 2020-01-11

\begin{verbatim}
(key) .bool_set:N = \boolean
\end{verbatim}

Defines \textit{(key)} to set \textit{(boolean)} to \textit{(value)} (which must be either \texttt{true} or \texttt{false}). If the \texttt{variable} does not exist, it will be created globally at the point that the key is set up.

\begin{verbatim}
(key) .bool_set_inverse:N = \boolean
\end{verbatim}

Defines \textit{(key)} to set \textit{(boolean)} to the logical inverse of \textit{(value)} (which must be either \texttt{true} or \texttt{false}). If the \texttt{boolean} does not exist, it will be created globally at the point that the key is set up.

\begin{verbatim}
(key) .choices:nn = { \choices } \{ \code \}
\end{verbatim}

Sets \textit{(key)} to act as a choice key. Each valid choice for \textit{(key)} must then be created, as discussed in section \texttt{3}.

\begin{verbatim}
(key) .clist_set:N = \comma list variable
\end{verbatim}

Defines \textit{(key)} to set \textit{(comma list variable)} to \textit{(value)}. Spaces around commas and empty items will be stripped. If the \texttt{variable} does not exist, it is created globally at the point that the key is set up.

\begin{verbatim}
(key) .code:n = \{ \code \}
\end{verbatim}

Stores the \texttt{(code)} for execution when \textit{(key)} is used. The \texttt{(code)} can include one parameter \texttt{(#1)}, which will be the \texttt{(value)} given for the \textit{(key)}.

\begin{verbatim}
(key) .cs_set:Np = \{control sequence\} \{arg. spec.\}
\end{verbatim}

Defines \textit{(key)} to set \textit{(control sequence)} to have \textit{(arg. spec.)} and replacement text \textit{(value)}.
\keys_define:nn { mymodule }
{  
  key .code:n = Hello-#1,  
  key .default:n = World  
}
\keys_set:nn { mymodule }
{  
  key = Fred, % Prints 'Hello Fred'  
  key, % Prints 'Hello World'  
  key = , % Prints 'Hello '  
}

The default does not affect keys where values are required or forbidden. Thus a required value cannot be supplied by a default value, and giving a default value for a key which cannot take a value does not trigger an error.

\keys_define:nn { mymodule }
{  
  key .code:n = Hello-#1,  
  key .default:n = World  
}
\keys_set:nn { mymodule }
{  
  key = Fred, % Prints 'Hello Fred'  
  key, % Prints 'Hello World'  
  key = , % Prints 'Hello '  
}

The default does not affect keys where values are required or forbidden. Thus a required value cannot be supplied by a default value, and giving a default value for a key which cannot take a value does not trigger an error.

\keys_define:nn { mymodule }
{  
  key .code:n = Hello-#1,  
  key .default:n = World  
}
\keys_set:nn { mymodule }
{  
  key = Fred, % Prints 'Hello Fred'  
  key, % Prints 'Hello World'  
  key = , % Prints 'Hello '  
}

The default does not affect keys where values are required or forbidden. Thus a required value cannot be supplied by a default value, and giving a default value for a key which cannot take a value does not trigger an error.

\keys_define:nn { mymodule }
{  
  key .code:n = Hello-#1,  
  key .default:n = World  
}
\keys_set:nn { mymodule }
{  
  key = Fred, % Prints 'Hello Fred'  
  key, % Prints 'Hello World'  
  key = , % Prints 'Hello '  
}

The default does not affect keys where values are required or forbidden. Thus a required value cannot be supplied by a default value, and giving a default value for a key which cannot take a value does not trigger an error.

\keys_define:nn { mymodule }
{  
  key .code:n = Hello-#1,  
  key .default:n = World  
}
\keys_set:nn { mymodule }
{  
  key = Fred, % Prints 'Hello Fred'  
  key, % Prints 'Hello World'  
  key = , % Prints 'Hello '  
}

The default does not affect keys where values are required or forbidden. Thus a required value cannot be supplied by a default value, and giving a default value for a key which cannot take a value does not trigger an error.

\keys_define:nn { mymodule }
{  
  key .code:n = Hello-#1,  
  key .default:n = World  
}
\keys_set:nn { mymodule }
{  
  key = Fred, % Prints 'Hello Fred'  
  key, % Prints 'Hello World'  
  key = , % Prints 'Hello '  
}

The default does not affect keys where values are required or forbidden. Thus a required value cannot be supplied by a default value, and giving a default value for a key which cannot take a value does not trigger an error.
### .initial:n
Initialises the \langle key \rangle with the \langle value \rangle, equivalent to
\[
\keys_set:nn \{\langle module \rangle\} \{ \langle key \rangle = \langle value \rangle \}
\]

### .int_set:N
\langle key \rangle .int_set:N = \langle integer \rangle
Defines \langle key \rangle to set \langle integer \rangle to \langle value \rangle (which must be an integer expression). If the variable does not exist, it is created globally at the point that the key is set up. The key will require a value at point-of-use unless a default is set.

### .meta:n
\langle key \rangle .meta:n = \{\langle keyval list \rangle\}
Makes \langle key \rangle a meta-key, which will set \langle keyval list \rangle in one go. The \langle keyval list \rangle can refer as \#1 to the value given at the time the \langle key \rangle is used (or, if no value is given, the \langle key \rangle’s default value).

### .meta:nn
\langle key \rangle .meta:nn = \{\langle path \rangle\} \{\langle keyval list \rangle\}
Makes \langle key \rangle a meta-key, which will set \langle keyval list \rangle in one go using the \langle path \rangle in place of the current one. The \langle keyval list \rangle can refer as \#1 to the value given at the time the \langle key \rangle is used (or, if no value is given, the \langle key \rangle’s default value).

### .multichoice:
\langle key \rangle .multichoice:
Sets \langle key \rangle to act as a multiple choice key. Each valid choice for \langle key \rangle must then be created, as discussed in section 3.

### .multichoices:nn
\langle key \rangle .multichoices:nn \{\langle choices \rangle\} \{\langle code \rangle\}
Sets \langle key \rangle to act as a multiple choice key, and defines a series \langle choices \rangle which are implemented using the \langle code \rangle. Inside \langle code \rangle, \l_keys_choice_tl will be the name of the choice made, and \l_keys_choice_int will be the position of the choice in the list of \langle choices \rangle (indexed from 1). Choices are discussed in detail in section 3.

### .muskip_set:N
\langle key \rangle .muskip_set:N = \langle muskip \rangle
Defines \langle key \rangle to set \langle muskip \rangle to \langle value \rangle (which must be a muskip expression). If the variable does not exist, it is created globally at the point that the key is set up. The key will require a value at point-of-use unless a default is set.

### .prop_put:N
\langle key \rangle .prop_put:N = \langle property list \rangle
Defines \langle key \rangle to put the \langle value \rangle onto the \langle property list \rangle stored under the \langle key \rangle. If the variable does not exist, it is created globally at the point that the key is set up.
2 Sub-dividing keys

When creating large numbers of keys, it may be desirable to divide them into several sub-groups for a given module. This can be achieved either by adding a sub-division to the module name:

\keys_define:nn { mymodule / subgroup }
{ key .code:n = code }

or to the key name:

\keys_define:nn { mymodule }
{ subgroup / key .code:n = code }

As illustrated, the best choice of token for sub-dividing keys in this way is /. This is because of the method that is used to represent keys internally. Both of the above code fragments set the same key, which has full name mymodule/subgroup/key.

As illustrated in the next section, this subdivision is particularly relevant to making multiple choices.
3 Choice and multiple choice keys

The \texttt{\textbackslash keys} system supports two types of choice key, in which a series of pre-defined input values are linked to varying implementations. Choice keys are usually created so that the various values are mutually-exclusive: only one can apply at any one time. “Multiple” choice keys are also supported: these allow a selection of values to be chosen at the same time.

Mutually-exclusive choices are created by setting the \texttt{.choice:} property:

\begin{verbatim}
\keys_define:nn { mymodule }
  { key .choice: }
\end{verbatim}

For keys which are set up as choices, the valid choices are generated by creating sub-keys of the choice key. This can be carried out in two ways.

In many cases, choices execute similar code which is dependant only on the name of the choice or the position of the choice in the list of all possibilities. Here, the keys can share the same code, and can be rapidly created using the \texttt{.choices:nn} property.

\begin{verbatim}
\keys_define:nn { mymodule }
  { key .choices:nn =
    { choice-a, choice-b, choice-c }
    { You\-gave\-choice\-'\textbackslash tl\_use:N \l\_keys\_choice_tl',~
      which\-is\-in\-position\-\textbackslash int\_use:N \l\_keys\_choice_int \textbackslash c\_space\_tl
       in\-the\-list. }
  }
\end{verbatim}

The index \texttt{\l\_keys\_choice_int} in the list of choices starts at 1.

\begin{verbatim}
\keys_define:nn { mymodule }
  { key .choice:,
    key / choice-a .code:n = code-a,
    key / choice-b .code:n = code-b,
    key / choice-c .code:n = code-c,
  }
\end{verbatim}

It is possible to mix the two methods, but manually-created choices should not use \texttt{\l\_keys\_choice_tl} or \texttt{\l\_keys\_choice_int}. These variables do not have defined
behaviour when used outside of code created using \texttt{.choices:nn} (\textit{i.e.} anything might happen).

It is possible to allow choice keys to take values which have not previously been defined by adding code for the special \texttt{unknown} choice. The general behavior of the \texttt{unknown} key is described in Section 5. A typical example in the case of a choice would be to issue a custom error message:

\begin{verbatim}
\keys_define:nn { mymodule }
{ key .choice:, key / choice-a .code:n = code-a, key / choice-b .code:n = code-b, key / choice-c .code:n = code-c, key / unknown .code:n = \msg_error:nxxx { mymodule } { unknown-choice }
{ key } % Name of choice key { choice-a , choice-b , choice-c } % Valid choices { \exp_not:n {#1} } % Invalid choice given % % }
\end{verbatim}

Multiple choices are created in a very similar manner to mutually-exclusive choices, using the properties \texttt{.multichoice:} and \texttt{.multichoices:nn}. As with mutually exclusive choices, multiple choices are defined as sub-keys. Thus both

\begin{verbatim}
\keys_define:nn { mymodule }
{ key .multichoices:nn = { choice-a, choice-b, choice-c } %
{ You-gave-choice-`\tl_use:N \l_keys_choice_tl',-
which-is-in-position-
\int_use:N \l_keys_choice_int \c_space_tl
in-the-list. }
}
\end{verbatim}

and

\begin{verbatim}
\keys_define:nn { mymodule }
{ key .multichoice:, key / choice-a .code:n = code-a, key / choice-b .code:n = code-b, key / choice-c .code:n = code-c, }
\end{verbatim}

are valid. When a multiple choice key is set
each choice is applied in turn, equivalent to a clist mapping or to applying each value individually:

```
\keys_set:nn { mymodule }
{  
key = a ,
key = b ,
key = c ,
}
```

Thus each separate choice will have passed to it the \l_keys_choice_tl and \l_keys_choice_int in exactly the same way as described for .choices:nn.

## 4 Setting keys

```
\keys_set:nn ⟨module⟩ (⟨keyval list⟩)
```

Parses the ⟨keyval list⟩, and sets those keys which are defined for ⟨module⟩. The behaviour on finding an unknown key can be set by defining a special unknown key: this is illustrated later.

For each key processed, information of the full path of the key, the name of the key and the value of the key is available within three token list variables. These may be used within the code of the key.

- **Path**: everything after the =, which may be empty if no value was given. This is stored in \l_keys_value_tl, and is not processed in any way by \keys_set:nn.
- **Name**: the part of the path after the last /, and thus is not unique. In the preceding examples, both keys have name key-a despite having different paths. This information is stored in \l_keys_key_str.

## 5 Handling of unknown keys

If a key has not previously been defined (is unknown), \keys_set:nn looks for a special unknown key for the same module, and if this is not defined raises an error indicating that
the key name was unknown. This mechanism can be used for example to issue custom error texts.

```
\keys_define:nn { mymodule }
{
  unknown .code:n =
    You-tried-to-set-key-'\l_keys_key_str'-to-'#1'.
}
```

These functions set keys which are known for the \textit{module}, and simply ignore other keys. The \texttt{\keys_set_known:nn} function parses the \textit{keyval list}, and sets those keys which are defined for \textit{module}. Any keys which are unknown are not processed further by the parser. In addition, \texttt{\keys_set_known:nnN} stores the key–value pairs in the \textit{tl} in comma-separated form \textit{\textit{i.e.} an edited version of the \textit{keyval list}}. When a \textit{root} is given \texttt{\keys_set_known:nnnN}, the key–value entries are returned relative to this point in the key tree. When it is absent, only the key name and value are provided. The correct list is returned by nested calls.

6 Selective key setting

In some cases it may be useful to be able to select only some keys for setting, even though these keys have the same path. For example, with a set of keys defined using

```
\keys_define:nn { mymodule }
{
  key-one .code:n = \my_func:n {#1} ,
  key-two .tl_set:N = \l_my_a_tl ,
  key-three .tl_set:N = \l_my_b_tl ,
  key-four .fp_set:N = \l_my_a_fp ,
}
```

the use of \texttt{\keys_set:nn} attempts to set all four keys. However, in some contexts it may only be sensible to set some keys, or to control the order of setting. To do this, keys may be assigned to \texttt{groups}: arbitrary sets which are independent of the key tree. Thus modifying the example to read

```
\keys_define:nn { mymodule }
{
  key-one .code:n = \my_func:n {#1} ,
  key-one .groups:n = \{ first \} ,
  key-two .tl_set:N = \l_my_a_tl ,
```
key-two .groups:n = { first },
key-three .tl_set:N = \l_my_b_tl,
key-three .groups:n = { second },
key-four .fp_set:N = \l_my_a_fp,
}\}
assigns key-one and key-two to group first, key-three to group second, while key-four is not assigned to a group.

Selective key setting may be achieved either by selecting one or more groups to be made “active”, or by marking one or more groups to be ignored in key setting.

\keys_set_filter:nnn \keys_set_filter:nnnN \keys_set_filter:nnnnN Activates key filtering in an “opt-out” sense: keys assigned to any of the (groups) specified are ignored. The (groups) are given as a comma-separated list. Unknown keys are not assigned to any group and are thus always set. The key–value pairs for each key which is filtered out are stored in the (tl) in a comma-separated form (i.e. an edited version of the (keyval list)). The \keys_set_filter:nnn version skips this stage.

Use of \keys_set_filter:nnnN can be nested, with the correct residual (keyval list) returned at each stage. In the version which takes a (root) argument, the key list is returned relative to that point in the key tree. In the cases without a (root) argument, only the key names and values are returned.

\keys_set_groups:nnn \keys_set_groups:nnnn \keys_set_groups:nnnnN \keys_set_groups:nnnnNN Activates key filtering in an “opt-in” sense: only keys assigned to one or more of the (groups) specified are set. The (groups) are given as a comma-separated list. Unknown keys are not assigned to any group and are thus never set.

7 Utility functions for keys

\keys_if_exist_p:nn * \keys_if_exist:nnTF * Tests if the (key) exists for (module), i.e. if any code has been defined for (key).
Tests if the \texttt{(choice)} is defined for the \texttt{(key)} within the \texttt{(module)}, i.e. if any code has been defined for \texttt{(key)/(choice)}. The test is \texttt{false} if the \texttt{(key)} itself is not defined.

Displays in the terminal the information associated to the \texttt{(key)} for a \texttt{(module)}, including the function which is used to actually implement it.

Writes in the log file the information associated to the \texttt{(key)} for a \texttt{(module)}. See also \texttt{\keys_show:nn} which displays the result in the terminal.

8 Low-level interface for parsing key–val lists

To re-cap from earlier, a key–value list is input of the form

\begin{verbatim}
KeyOne = ValueOne ,
KeyTwo = ValueTwo ,
KeyThree
\end{verbatim}

where each key–value pair is separated by a comma from the rest of the list, and each key–value pair does not necessarily contain an equals sign or a value! Processing this type of input correctly requires a number of careful steps, to correctly account for braces, spaces and the category codes of separators.

While the functions described earlier are used as a high-level interface for processing such input, in special circumstances you may wish to use a lower-level approach. The low-level parsing system converts a \texttt{(key–value list)} into \texttt{(keys)} and associated \texttt{(values)}. After the parsing phase is completed, the resulting keys and values (or keys alone) are available for further processing. This processing is not carried out by the low-level parser itself, and so the parser requires the names of two functions along with the key–value list. One function is needed to process key–value pairs (it receives two arguments), and a second function is required for keys given without any value (it is called with a single argument).

The parser does not double \texttt{#} tokens or expand any input. Active tokens = and \texttt{,} appearing at the outer level of braces are converted to category “other” (12) so that the parser does not “miss” any due to category code changes. Spaces are removed from the ends of the keys and values. Keys and values which are given in braces have exactly one set removed (after space trimming), thus

\begin{verbatim}
key = {value here},
\end{verbatim}

and

\begin{verbatim}
key = value here,
\end{verbatim}
are treated identically.

\keyval_parse:NNn * \keyval_parse:NNn \{function\1 \function\2 \{\key–value list\}\}

Parses the \{key–value list\} into a series of \{keys\} and associated \{values\}, or keys alone (if no \{value\} was given). \{function\1\} should take one argument, while \{function\2\} should absorb two arguments. After \keyval_parse:NNn has parsed the \{key–value list\}, \{function\1\} is used to process keys given with no value and \{function\2\} is used to process keys given with a value. The order of the \{keys\} in the \{key–value list\} is preserved. Thus

\keyval_parse:NNn \function:n \function:nn
\{ key1 = value1 , key2 = value2 , key3 = , key4 \}

is converted into an input stream

\function:nn \{ key1 \} \{ value1 \}
\function:nn \{ key2 \} \{ value2 \}
\function:nn \{ key3 \} \{ \}
\function:n \{ key4 \}

Note that there is a difference between an empty value (an equals sign followed by nothing) and a missing value (no equals sign at all). Spaces are trimmed from the ends of the \{key\} and \{value\}, then one outer set of braces is removed from the \{key\} and \{value\} as part of the processing.

\TeX\ Hackers note: The result is returned within \exp_not:n, which means that the converted input stream does not expand further when appearing in an x-type or e-type argument expansion.
Part XXII
The \texttt{l3intarray} package: fast global integer arrays

1 \texttt{l3intarray} documentation

For applications requiring heavy use of integers, this module provides arrays which can be accessed in constant time (contrast \texttt{l3seq}, where access time is linear). These arrays have several important features:

- The size of the array is fixed and must be given at point of initialisation
- The absolute value of each entry has maximum $2^{30} - 1$ (\emph{i.e.} one power lower than the usual \texttt{\maxint} ceiling of $2^{31} - 1$)

The use of \texttt{intarray} data is therefore recommended for cases where the need for fast access is of paramount importance.

\begin{verbatim}
\intarray_new:Nn \intarray_new:cn
\hline
\verb|\intarray_new:Nn\{intarray\ var\}\{(size)\}|
\hline
\verb|\intarray_new:cn|
\verb|New: 2018-03-29|
\end{verbatim}

Evaluates the integer expression \texttt{(size)} and allocates an \texttt{integer array variable} with that number of (zero) entries. The variable name should start with \texttt{\g_} because assignments are always global.

\begin{verbatim}
\intarray_count:N \intarray_count:c 
\hline
\verb|\intarray_count:N\{intarray\ var\}|
\verb|\intarray_count:c|
\hline
\verb|\intarray_count:N\ \\intarray_count:c|
\verb|New: 2018-03-29|
\end{verbatim}

Expands to the number of entries in the \texttt{integer array variable}. Contrarily to \texttt{\seq_-count:N} this is performed in constant time.

\begin{verbatim}
\intarray_gset:Nnn \intarray_gset:cn
\hline
\verb|\intarray_gset:Nnn\{intarray\ var\}\{\{position\}\}\{\{value\}\}|
\hline
\verb|\intarray_gset:cn|
\verb|New: 2018-03-29|
\end{verbatim}

Stores the result of evaluating the integer expression \texttt{(value)} into the \texttt{integer array variable} at the (integer expression) \texttt{(position)}. If the \texttt{(position)} is not between 1 and the \texttt{\intarray_count:N}, or the \texttt{(value)}’s absolute value is bigger than $2^{30} - 1$, an error occurs. Assignments are always global.

\begin{verbatim}
\intarray_const_from_clist:Nn \intarray_const_from_clist:cn
\hline
\verb|\intarray_const_from_clist:Nn\{intarray\ var\}\\{intexpr\ clist\}|
\verb|\intarray_const_from_clist:cn|
\hline
\verb|\intarray_const_from_clist:Nn\ \\intarray_const_from_clist:cn|
\verb|New: 2018-05-04|
\end{verbatim}

Creates a new constant \texttt{integer array variable} or raises an error if the name is already taken. The \texttt{integer array variable} is set (globally) to contain as its items the results of evaluating each \texttt{integer expression} in the \texttt{comma list}.

\begin{verbatim}
\intarray_gzero:N \intarray_gzero:cn
\hline
\verb|\intarray_gzero:N\{intarray\ var\}|
\verb|\intarray_gzero:cn|
\hline
\verb|\intarray_gzero:N\ \\intarray_gzero:cn|
\verb|New: 2018-05-04|
\end{verbatim}

Sets all entries of the \texttt{integer array variable} to zero. Assignments are always global.
\texttt{\textbackslash intarray\_item:Nm} \texttt{(intarray var) \{(position)\}}

Expands to the integer entry stored at the (integer expression) \textit{(position)} in the \texttt{(integer array variable)}. If the \textit{(position)} is not between 1 and the \texttt{\textbackslash intarray\_count:N}, an error occurs.

\texttt{\textbackslash intarray\_rand\_item:Nm} \texttt{(intarray var)}

Selects a pseudo-random item of the \texttt{(integer array)}. If the \texttt{(integer array)} is empty, produce an error.

\texttt{\textbackslash intarray\_show:Nm} \texttt{(intarray var)}
\texttt{\textbackslash intarray\_log:Nm} \texttt{(intarray var)}

Displays the items in the \texttt{(integer array variable)} in the terminal or writes them in the log file.

1.1 Implementation notes

It is a wrapper around the \texttt{fontdimen} primitive, used to store arrays of integers (with a restricted range: absolute value at most $2^{30} - 1$). In contrast to \texttt{l3seq} sequences the access to individual entries is done in constant time rather than linear time, but only integers can be stored. More precisely, the primitive \texttt{fontdimen} stores dimensions but the \texttt{l3intarray} package transparently converts these from/to integers. Assignments are always global.

While Lua\TeX{}’s memory is extensible, other engines can “only” deal with a bit less than $4 \times 10^6$ entries in all \texttt{fontdimen} arrays combined (with default \TeX{}Live settings).
Part XXIII

The l3fp package: Floating points

A decimal floating point number is one which is stored as a significand and a separate exponent. The module implements expandably a wide set of arithmetic, trigonometric, and other operations on decimal floating point numbers, to be used within floating point expressions. Floating point expressions support the following operations with their usual precedence.

- Basic arithmetic: addition $x + y$, subtraction $x - y$, multiplication $x \times y$, division $x/y$, square root $\sqrt{x}$, and parentheses.
- Comparison operators: $x < y$, $x \leq y$, $x > y$, $x < y$ etc.
- Boolean logic: sign $\text{sign}(x)$, negation $!x$, conjunction $x \& \& y$, disjunction $x \| y$, ternary operator $x ? y : z$.
- Exponentials: $\exp x$, $\ln x$, $x^y$, $\log_b x$.
- Integer factorial: $\text{fact} x$.
- Trigonometry: $\sin x$, $\cos x$, $\tan x$, $\cot x$, $\sec x$, $\csc x$ expecting their arguments in radians, and $\sin d x$, $\cos d x$, $\tan d x$, $\cot d x$, $\sec d x$, $\csc d x$ expecting their arguments in degrees.
- Inverse trigonometric functions: $\arcsin x$, $\arccos x$, $\arctan x$, $\arc cot x$, $\arcsec x$, $\arccsc x$ giving a result in radians, and $\arcsind x$, $\arccosd x$, $\arctand x$, $\arc cotd x$, $\arcsecd x$, $\arccscd x$ giving a result in degrees.
- Hyperbolic functions and their inverse functions: $\sinh x$, $\cosh x$, $\tanh x$, $\coth x$, $\sech x$, $\csch x$, $\text{arcsinh} x$, $\text{arccosh} x$, $\text{arctanh} x$, $\text{arccoth} x$, $\text{arcsech} x$, $\text{arccsch} x$.
- Extrema: $\max(x_1, x_2, \ldots)$, $\min(x_1, x_2, \ldots)$, $\text{abs}(x)$.
- Rounding functions, controlled by two optional values, $n$ (number of places, 0 by default) and $t$ (behavior on a tie, NaN by default):
  - $\text{trunc}(x, n)$ rounds towards zero,
  - $\text{floor}(x, n)$ rounds towards $-\infty$,
  - $\text{ceil}(x, n)$ rounds towards $+\infty$,
  - $\text{round}(x, n, t)$ rounds to the closest value, with ties rounded to an even value by default, towards zero if $t = 0$, towards $+\infty$ if $t > 0$ and towards $-\infty$ if $t < 0$.
- And (not yet) modulo, and “quantize”.
- Random numbers: $\text{rand}()$, $\text{randint}(m, n)$.
- Constants: $\pi$, $\text{deg}$ (one degree in radians).
- Dimensions, automatically expressed in points, e.g., $\text{pc}$ is 12.
Automatic conversion (no need for \(\text{\texttt{ type \_\texttt{use}:N} \)} of integer, dimension, and skip variables to floating point numbers, expressing dimensions in points and ignoring the stretch and shrink components of skips.

- Tuples: \((x_1,\ldots,x_n)\) that can be stored in variables, added together, multiplied or divided by a floating point number, and nested.

Floating point numbers can be given either explicitly (in a form such as \(1.234e^{-34}\), or \(-.0001\)), or as a stored floating point variable, which is automatically replaced by its current value. A “floating point” is a floating point number or a tuple thereof. See section 9.1 for a description of what a floating point is, section 9.2 for details about how an expression is parsed, and section 9.3 to know what the various operations do. Some operations may raise exceptions (error messages), described in section 7.

An example of use could be the following.

\[
\LaTeX{} \text{can now compute: } \frac{\sin (3.5)}{2} + 2 \cdot 10^{-3} = \text{\texttt{\ExplSyntaxOn \fp_to_decimal:n {sin(3.5)/2 + 2e-3}}}.
\]

The operation \texttt{round} can be used to limit the result’s precision. Adding +0 avoids the possibly undesirable output -0, replacing it by +0. However, the l3fp module is mostly meant as an underlying tool for higher-level commands. For example, one could provide a function to typeset nicely the result of floating point computations.

\[
\documentclass{article}
\usepackage{xparse, siunitx}
\ExplSyntaxOn
\NewDocumentCommand { \calcnum } { m }{ \num { \fp_to_scientific:n {#1} } }
\ExplSyntaxOff
\begin{document}
calcnum { 2 \pi \ast \sin\left( 2.3 ^ 5 \right) }
\end{document}
\]

See the documentation of siunitx for various options of \texttt{\num}.

\section{Creating and initialising floating point variables}

\begin{verbatim}
\fp_new:N \fp_new:c
Updated: 2012-05-08
\fp_new:N \fp_new:c
\fp_zero:N \fp_zero:c \fp_gzero:N \fp_gzero:c
Updated: 2012-05-08
\end{verbatim}

\begin{itemize}
\item \texttt{\fp_new:N \texttt{\texttt{(fp var)}}} creates a new \texttt{(fp var)} or raises an error if the name is already taken. The declaration is global. The \texttt{(fp var)} is initially +0.
\item \texttt{\fp_const:Nn \texttt{(fp var)} \texttt{\{floating point expression\}}} creates a new constant \texttt{(fp var)} or raises an error if the name is already taken. The \texttt{(fp var)} is set globally equal to the result of evaluating the \texttt{(floating point expression)}.
\item \texttt{\fp_zero:N \texttt{(fp var)}} sets the \texttt{(fp var)} to +0.
\end{itemize}
Ensures that the \textit{fp var} exists globally by applying \texttt{\_\_\_\_\_\_\_new:N} if necessary, then applies \texttt{\_\_\_\_\_\_gzero:N} to leave the \textit{fp var} set to +0.

2 Setting floating point variables

\texttt{\_\_\_\_\_\_\_set:Nn} \textit{fp var} \{\textit{floating point expression}\}

Sets \textit{fp var} equal to the result of computing the \{\textit{floating point expression}\}.

\texttt{\_\_\_\_\_\_\_set_eq:NN} \textit{fp var}_1 \textit{fp var}_2

Sets the floating point variable \textit{fp var}_1 equal to the current value of \textit{fp var}_2.

\texttt{\_\_\_\_\_\_\_add:Nn} \textit{fp var} \{\textit{floating point expression}\}

Adds the result of computing the \{\textit{floating point expression}\} to the \textit{fp var}. This also applies if \textit{fp var} and \{\textit{floating point expression}\} evaluate to tuples of the same size.

\texttt{\_\_\_\_\_\_\_sub:Nn} \textit{fp var} \{\textit{floating point expression}\}

Subtracts the result of computing the \{\textit{floating point expression}\} from the \textit{fp var}. This also applies if \textit{fp var} and \{\textit{floating point expression}\} evaluate to tuples of the same size.

3 Using floating points

\texttt{\_\_\_\_\_\_\_eval:n} \{\textit{floating point expression}\}

Evaluates the \{\textit{floating point expression}\} and expresses the result as a decimal number with no exponent. Leading or trailing zeros may be inserted to compensate for the exponent. Non-significant trailing zeros are trimmed, and integers are expressed without a decimal separator. The values ±∞ and NaN trigger an “invalid operation” exception. For a tuple, each item is converted using \texttt{\_\_\_\_\_\_\_eval:n} and they are combined as \((\textit{fp var}_1, \textit{fp var}_2, \ldots, \textit{fp var}_n)\) if \(n > 1\) and \((\textit{fp var}_1)\) or \(()\) for fewer items. This function is identical to \texttt{\_\_\_\_\_\_\_to_decimal:n}.
\fp_sign:n \{\langle \text{fpexpr} \rangle \}\}

Evaluates the \langle \text{fpexpr} \rangle and leaves its sign in the input stream using \fp_eval:n \{\text{sign}(\langle \text{result} \rangle)\}: +1 for positive numbers and for \(+\infty\), −1 for negative numbers and for \(−\infty\), ±0 for ±0. If the operand is a tuple or is \texttt{NaN}, then “invalid operation” occurs and the result is 0.

\fp_to_decimal:N \{\langle \text{fp var} \rangle \}\}
\fp_to_decimal:c \{\langle \text{floating point expression} \rangle \}\}
\fp_to_decimal:n \{\langle \text{floating point expression} \rangle \}\}

Evaluates the \langle \text{floating point expression} \rangle and expresses the result as a decimal number with no exponent. Leading or trailing zeros may be inserted to compensate for the exponent. Non-significant trailing zeros are trimmed, and integers are expressed without a decimal separator. The values \(\pm\infty\) and \texttt{NaN} trigger an “invalid operation” exception. For a tuple, each item is converted using \fp_to_decimal:n and they are combined as \((\langle \text{fp}_1 \rangle, \langle \text{fp}_2 \rangle, \ldots, \langle \text{fp}_n \rangle)\) if \(n > 1\) and \((\langle \text{fp}_1 \rangle, \ldots)\) or () for fewer items.

\fp_to_dim:N \{\langle \text{fp var} \rangle \}\}
\fp_to_dim:c \{\langle \text{floating point expression} \rangle \}\}
\fp_to_dim:n \{\langle \text{floating point expression} \rangle \}\}

Evaluates the \langle \text{floating point expression} \rangle and expresses the result as a dimension (in \texttt{pt}) suitable for use in dimension expressions. The output is identical to \fp_to_decimal:n, with an additional trailing \texttt{pt} (both letter tokens). In particular, the result may be outside the range \([-2^{14} + 2^{-17}, 2^{14} - 2^{-17}]\) of valid \TeX{} dimensions, leading to overflow errors if used as a dimension. Tuples, as well as the values \(\pm\infty\) and \texttt{NaN}, trigger an “invalid operation” exception.

\fp_to_int:N \{\langle \text{fp var} \rangle \}\}
\fp_to_int:c \{\langle \text{floating point expression} \rangle \}\}
\fp_to_int:n \{\langle \text{floating point expression} \rangle \}\}

Evaluates the \langle \text{floating point expression} \rangle, and rounds the result to the closest integer, rounding exact ties to an even integer. The result may be outside the range \([-2^{31} + 1, 2^{31} - 1]\) of valid \TeX{} integers, leading to overflow errors if used in an integer expression. Tuples, as well as the values \(\pm\infty\) and \texttt{NaN}, trigger an “invalid operation” exception.

\fp_to_scientific:N \{\langle \text{fp var} \rangle \}\}
\fp_to_scientific:c \{\langle \text{floating point expression} \rangle \}\}
\fp_to_scientific:n \{\langle \text{floating point expression} \rangle \}\}

Evaluates the \langle \text{floating point expression} \rangle and expresses the result in scientific notation:

\langle \text{optional }-\rangle \langle \text{digit} \rangle . \langle 15 \text{ digits} \rangle \langle \text{optional sign} \rangle \langle \text{e} \rangle \langle \text{optional sign} \rangle \langle \text{exponent} \rangle

The leading \langle \text{digit} \rangle is non-zero except in the case of ±0. The values \(\pm\infty\) and \texttt{NaN} trigger an “invalid operation” exception. Normal category codes apply: thus the \texttt{e} is category code 11 (a letter). For a tuple, each item is converted using \fp_to_scientific:n and they are combined as \((\langle \text{fp}_1 \rangle, \ldots, \langle \text{fp}_n \rangle)\) if \(n > 1\) and \((\langle \text{fp}_1 \rangle, \ldots)\) or () for fewer items.
4 Floating point conditionals

\begin{itemize}
  \item \texttt{\textbackslash fp\_compare\_nNn}\{\texttt{fpexpr}_1\}\{\texttt{relation}\}\{\texttt{fpexpr}_2\}\{\texttt{true\ code}\}\{\texttt{false\ code}\}
  \item \texttt{\textbackslash fp\_compare\_nNnTF}\{\texttt{fpexpr}_1\}\{\texttt{relation}\}\{\texttt{fpexpr}_2\}\{\texttt{true\ code}\}\{\texttt{false\ code}\}
\end{itemize}

Compares the \texttt{\textbackslash fpexpr}_1 and the \texttt{\textbackslash fpexpr}_2, and returns \texttt{true} if the \texttt{\textbackslash relation} is obeyed. Two floating points \texttt{x} and \texttt{y} may obey four mutually exclusive relations: \texttt{x < y}, \texttt{x = y}, \texttt{x > y}, or \texttt{x?y} (“not ordered”). The last case occurs exactly if one or both operands is \texttt{NaN} or is a tuple, unless they are equal tuples. Note that a \texttt{NaN} is distinct from any value, even another \texttt{NaN}, hence \texttt{x = x} is not true for a \texttt{NaN}. To test if a value is \texttt{NaN}, compare it to an arbitrary number with the “not ordered” relation.

\begin{verbatim}
\texttt{\textbackslash fp\_compare\_nNnTF \{<value> \} ? 0 }
\{ } % <value> is nan
\{ } % <value> is not nan
\end{verbatim}

Tuples are equal if they have the same number of items and items compare equal (in particular there must be no \texttt{NaN}). At present any other comparison with tuples yields \texttt{?} (not ordered). This is experimental.

This function is less flexible than \texttt{\textbackslash fp\_compare\_nTF} but slightly faster. It is provided for consistency with \texttt{\textbackslash int\_compare\_nNnTF} and \texttt{\textbackslash dim\_compare\_nNnTF}.
Evaluates the floating point expressions as described for \fp_eval:n and compares consecutive result using the corresponding relation, namely it compares \intexpr{1} and \intexpr{2} using the \relation{1}, then \intexpr{2} and \intexpr{3} using the \relation{2}, until finally comparing \intexpr{N} and \intexpr{N+1} using the \relation{N}. The test yields true if all comparisons are true. Each floating point expression is evaluated only once. Contrarily to \int_compare:nTF, all floating point expressions are computed, even if one comparison is false. Two floating points \(x\) and \(y\) may obey four mutually exclusive relations: \(x < y\), \(x = y\), \(x > y\), or \(x \not\sim y\) (“not ordered”). The last case occurs exactly if one or both operands is \texttt{NaN} or is a tuple, unless they are equal tuples. Each \relation{\(\cdots\)} can be any (non-empty) combination of \(<\), \(=\), \(>\), and \(?\), plus an optional leading \(!\) (which negates the \relation). with the restriction that the \relation may not start with \(?\), as this symbol has a different meaning (in combination with \:) within floating point expressions. The comparison \(x \rel y\) is then true if the \relation does not start with \(!\) and the actual relation (\(<\), \(=\), \(>\), or \(?\)) between \(x\) and \(y\) appears within the \relation, or on the contrary if the \relation starts with \(!\) and the relation between \(x\) and \(y\) does not appear within the \relation. Common choices of \relation include \(\geq\) (greater or equal), \(!=\) (not equal), \(!\?\) or \(\leftrightarrow\) (comparable).

This function is more flexible than \fp_compare:nNnTF and only slightly slower.

## 5 Floating point expression loops

\fp_do_until:nNnn \{ \fpexpr \relation \code \}

Places the \(\code\) in the input stream for \TeX{} to process, and then evaluates the relationship between the two \(\texttt{floating point expressions}\) as described for \fp_compare:nNnTF. If the test is false then the \(\code\) is inserted into the input stream again and a loop occurs until the \relation is true.

\fp_do_while:nNnn \{ \fpexpr \relation \code \}

Places the \(\code\) in the input stream for \TeX{} to process, and then evaluates the relationship between the two \(\texttt{floating point expressions}\) as described for \fp_compare:nNnTF. If the test is true then the \(\code\) is inserted into the input stream again and a loop occurs until the \relation is false.
\texttt{\textbackslash fp\textunderscore until\textunderscore do:nNnn} \* \texttt{\textbackslash fp\textunderscore until\textunderscore do:nNnn \{fpexpr\1\} \{relation\} \{fpexpr\2\} \{(code)\}}

Evaluates the relationship between the two \textit{(floating point expressions)} as described for \texttt{\textbackslash fp\textunderscore compare:nNnTF}, and then places the \textit{(code)} in the input stream if the \textit{(relation)} is \textit{false}. After the \textit{(code)} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \textit{true}.\[\text{New: 2012-08-16}\]

\texttt{\textbackslash fp\textunderscore while\textunderscore do:nNnn} \* \texttt{\textbackslash fp\textunderscore while\textunderscore do:nNnn \{fpexpr\1\} \{relation\} \{fpexpr\2\} \{(code)\}}

Evaluates the relationship between the two \textit{(floating point expressions)} as described for \texttt{\textbackslash fp\textunderscore compare:nNnTF}, and then places the \textit{(code)} in the input stream if the \textit{(relation)} is \textit{true}. After the \textit{(code)} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \textit{false}.\[\text{New: 2012-08-16}\]

\texttt{\textbackslash fp\textunderscore do\textunderscore until:nn} \* \texttt{\textbackslash fp\textunderscore do\textunderscore until:nn \{fpexpr\1\} \{relation\} \{fpexpr\2\} \{(code)\}}

Places the \textit{(code)} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \textit{(floating point expressions)} as described for \texttt{\textbackslash fp\textunderscore compare:nNnTF}. If the test is \textit{false} then the \textit{(code)} is inserted into the input stream again and a loop occurs until the \textit{(relation)} is \textit{true}.\[\text{New: 2012-08-16} \quad \text{Updated: 2013-12-14}\]

\texttt{\textbackslash fp\textunderscore do\textunderscore while:nn} \* \texttt{\textbackslash fp\textunderscore do\textunderscore while:nn \{fpexpr\1\} \{relation\} \{fpexpr\2\} \{(code)\}}

Places the \textit{(code)} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \textit{(floating point expressions)} as described for \texttt{\textbackslash fp\textunderscore compare:nNnTF}. If the test is \textit{true} then the \textit{(code)} is inserted into the input stream again and a loop occurs until the \textit{(relation)} is \textit{false}.\[\text{New: 2012-08-16} \quad \text{Updated: 2013-12-14}\]

\texttt{\textbackslash fp\textunderscore until\textunderscore do:nn} \* \texttt{\textbackslash fp\textunderscore until\textunderscore do:nn \{fpexpr\1\} \{relation\} \{fpexpr\2\} \{(code)\}}

Evaluates the relationship between the two \textit{(floating point expressions)} as described for \texttt{\textbackslash fp\textunderscore compare:nNnTF}, and then places the \textit{(code)} in the input stream if the \textit{(relation)} is \textit{false}. After the \textit{(code)} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \textit{true}.\[\text{New: 2012-08-16} \quad \text{Updated: 2013-12-14}\]

\texttt{\textbackslash fp\textunderscore while\textunderscore do:nn} \* \texttt{\textbackslash fp\textunderscore while\textunderscore do:nn \{fpexpr\1\} \{relation\} \{fpexpr\2\} \{(code)\}}

Evaluates the relationship between the two \textit{(floating point expressions)} as described for \texttt{\textbackslash fp\textunderscore compare:nNnTF}, and then places the \textit{(code)} in the input stream if the \textit{(relation)} is \textit{true}. After the \textit{(code)} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \textit{false}.\[\text{New: 2012-08-16} \quad \text{Updated: 2013-12-14}\]
\fp_step_function:nnnN \fp_step_function:nnnc

\textbf{New: 2016-11-21} \textbf{Updated: 2016-12-06}

\texttt{\fp_step_function:nnnN \{initial value\} \{step\} \{final value\} \{function\}}

This function first evaluates the \texttt{initial value}, \texttt{step} and \texttt{final value}, each of which should be a floating point expression evaluating to a floating point number, not a tuple. The \texttt{function} is then placed in front of each \texttt{value} from the \texttt{initial value} to the \texttt{final value} in turn (using \texttt{step} between each \texttt{value}). The \texttt{step} must be non-zero. If the \texttt{step} is positive, the loop stops when the \texttt{value} becomes larger than the \texttt{final value}. If the \texttt{step} is negative, the loop stops when the \texttt{value} becomes smaller than the \texttt{final value}. The \texttt{function} should absorb one numerical argument. For example

\begin{verbatim}
\cs_set:Npn \my_func:n #1 { \[I saw #1\] \quad }
\fp_step_function:nnnN { 1.0 } { 0.1 } { 1.5 } \my_func:n
\end{verbatim}

would print

\begin{verbatim}
[I saw 1.0] [I saw 1.1] [I saw 1.2] [I saw 1.3] [I saw 1.4] [I saw 1.5]
\end{verbatim}

\textbf{\TeX{} hackers note:} Due to rounding, it may happen that adding the \texttt{step} to the \texttt{value} does not change the \texttt{value}; such cases give an error, as they would otherwise lead to an infinite loop.

\fp_step_variable:nnnNn

\textbf{New: 2017-04-12}

\texttt{\fp_step_variable:nnnNn \{initial value\} \{step\} \{final value\} \{tl var\} \{code\}}

This function first evaluates the \texttt{initial value}, \texttt{step} and \texttt{final value}, all of which should be floating point expressions evaluating to a floating point number, not a tuple. Then for each \texttt{value} from the \texttt{initial value} to the \texttt{final value} in turn (using \texttt{step} between each \texttt{value}), the \texttt{code} is inserted into the input stream with \texttt{\#1} replaced by the current \texttt{value}. Thus the \texttt{code} should define a function of one argument (\texttt{\#1}).

\section{Some useful constants, and scratch variables}

\c_zero_fp \c_minus_zero_fp

\textbf{New: 2012-05-08}

Zero, with either sign.

\c_one_fp

\textbf{New: 2012-05-08}

One as an fp: useful for comparisons in some places.
Infinity, with either sign. These can be input directly in a floating point expression as inf and -inf.

The value of the base of the natural logarithm, \( e = \exp(1) \).

The value of \( \pi \). This can be input directly in a floating point expression as pi.

The value of 1° in radians. Multiply an angle given in degrees by this value to obtain a result in radians. Note that trigonometric functions expecting an argument in radians or in degrees are both available. Within floating point expressions, this can be accessed as deg.

Scratch floating points for local assignment. These are never used by the kernel code, and so are safe for use with any \( \text{LATEX3} \)-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

Scratch floating points for global assignment. These are never used by the kernel code, and so are safe for use with any \( \text{LATEX3} \)-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

7 Floating point exceptions

The functions defined in this section are experimental, and their functionality may be altered or removed altogether.

“Exceptions” may occur when performing some floating point operations, such as \( 0 / 0 \), or \( 10 ** \text{e9999} \). The relevant IEEE standard defines 5 types of exceptions, of which we implement 4.

- **Overflow** occurs whenever the result of an operation is too large to be represented as a normal floating point number. This results in \( \pm\infty \).
- **Underflow** occurs whenever the result of an operation is too close to 0 to be represented as a normal floating point number. This results in \( \pm0 \).
- **Invalid operation** occurs for operations with no defined outcome, for instance 0/0 or \( \sin(\infty) \), and results in a NaN. It also occurs for conversion functions whose target type does not have the appropriate infinite or NaN value (e.g., \( \text{fp_to_dim:n} \)).
- **Division by zero** occurs when dividing a non-zero number by 0, or when evaluating functions at poles, e.g., ln(0) or cot(0). This results in \( \pm\infty \).
\(\text{Inexact}\) occurs whenever the result of a computation is not exact, in other words, almost always. At the moment, this exception is entirely ignored in \(\LaTeX 3\).

To each exception we associate a “flag”: \texttt{fp\_overflow}, \texttt{fp\_underflow}, \texttt{fp\_invalid\_operation} and \texttt{fp\_division\_by\_zero}. The state of these flags can be tested and modified with commands from \texttt{l3flag}.

By default, the “invalid operation” exception triggers an (expandable) error, and raises the corresponding flag. Other exceptions raise the corresponding flag but do not trigger an error. The behaviour when an exception occurs can be modified (using \texttt{fp\_trap:nn}) to either produce an error and raise the flag, or only raise the flag, or do nothing at all.

\[
\texttt{fp\_trap:nn} \{\texttt{exception}\}\} \{\texttt{trap type}\} \\
\texttt{fp\_trap:nn} \{\texttt{exception}\}\} \{\texttt{trap type}\}
\]

All occurrences of the \{\texttt{exception}\} (\texttt{overflow}, \texttt{underflow}, \texttt{invalid\_operation} or \texttt{division\_by\_zero}) within the current group are treated as \{\texttt{trap type}\}, which can be

- \texttt{none}: the \{\texttt{exception}\} will be entirely ignored, and leave no trace;
- \texttt{flag}: the \{\texttt{exception}\} will turn the corresponding flag on when it occurs;
- \texttt{error}: additionally, the \{\texttt{exception}\} will halt the \TeX run and display some information about the current operation in the terminal.

This function is experimental, and may be altered or removed.

Flags denoting the occurrence of various floating-point exceptions.

8 Viewing floating points

\[
\texttt{fp\_show:N} \{\texttt{fp\ var}\} \\
\texttt{fp\_show:C} \{\texttt{floating point expression}\} \\
\texttt{fp\_show:N}
\]

Evaluates the \{\texttt{floating point expression}\} and displays the result in the terminal.

\[
\texttt{fp\_log:N} \{\texttt{fp\ var}\} \\
\texttt{fp\_log:C} \{\texttt{floating point expression}\} \\
\texttt{fp\_log:N}
\]

Evaluates the \{\texttt{floating point expression}\} and writes the result in the log file.
9 Floating point expressions

9.1 Input of floating point numbers

We support four types of floating point numbers:

- $\pm m \cdot 10^n$, a floating point number, with integer $1 \leq m \leq 10^{16}$, and $-10000 \leq n \leq 10000$;
- $\pm 0$, zero, with a given sign;
- $\pm \infty$, infinity, with a given sign;
- NaN, is “not a number”, and can be either quiet or signalling (not yet: this distinction is currently unsupported);

Normal floating point numbers are stored in base 10, with up to 16 significant figures.

On input, a normal floating point number consists of:

- $\langle \text{sign} \rangle$: a possibly empty string of + and - characters;
- $\langle \text{significand} \rangle$: a non-empty string of digits together with zero or one dot;
- $\langle \text{exponent} \rangle$ optionally: the character e, followed by a possibly empty string of + and - tokens, and a non-empty string of digits.

The sign of the resulting number is + if $\langle \text{sign} \rangle$ contains an even number of -, and - otherwise, hence, an empty $\langle \text{sign} \rangle$ denotes a non-negative input. The stored significand is obtained from $\langle \text{significand} \rangle$ by omitting the decimal separator and leading zeros, and rounding to 16 significant digits, filling with trailing zeros if necessary. In particular, the value stored is exact if the input $\langle \text{significand} \rangle$ has at most 16 digits. The stored $\langle \text{exponent} \rangle$ is obtained by combining the input $\langle \text{exponent} \rangle$ (0 if absent) with a shift depending on the position of the significand and the number of leading zeros.

A special case arises if the resulting $\langle \text{exponent} \rangle$ is either too large or too small for the floating point number to be represented. This results either in an overflow (the number is then replaced by $\pm \infty$), or an underflow (resulting in $\pm 0$).

The result is thus $\pm 0$ if and only if $\langle \text{significand} \rangle$ contains no non-zero digit (i.e., consists only in characters 0, and an optional period), or if there is an underflow. Note that a single dot is currently a valid floating point number, equal to +0, but that is not guaranteed to remain true.

The $\langle \text{significand} \rangle$ must be non-empty, so e1 and e-1 are not valid floating point numbers. Note that the latter could be mistaken with the difference of “e” and 1. To avoid confusions, the base of natural logarithms cannot be input as e and should be input as exp(1) or \c_e_fp.

Special numbers are input as follows:

- inf represents $+\infty$, and can be preceded by any $\langle \text{sign} \rangle$, yielding $\pm \infty$ as appropriate.
- nan represents a (quiet) non-number. It can be preceded by any sign, but that sign is ignored.
- Any unrecognizable string triggers an error, and produces a NaN.
- Note that commands such as \infty, \pi, or \sin do not work in floating point expressions. They may silently be interpreted as completely unexpected numbers, because integer constants (allowed in expressions) are commonly stored as mathematical characters.
9.2 Precedence of operators

We list here all the operations supported in floating point expressions, in order of decreasing precedence: operations listed earlier bind more tightly than operations listed below them.

- Function calls (\texttt{sin}, \texttt{ln}, \texttt{etc}).
- Binary \texttt{**} and \texttt{\^} (right associative).
- Unary \texttt{+}, \texttt{-}, \texttt{!}.
- Implicit multiplication by juxtaposition (2pi) when neither factor is in parentheses.
- Binary \texttt{*} and \texttt{/}, implicit multiplication by juxtaposition with parentheses (for instance 3(4+5)).
- Binary \texttt{+} and \texttt{-}.
- Comparisons \texttt{>=}, \texttt{!=}, \texttt{<?}, \texttt{etc}.
- Logical \texttt{and}, denoted by \texttt{&&}.
- Logical \texttt{or}, denoted by \texttt{||}.
- Ternary operator \texttt{?:} (right associative).
- Comma (to build tuples).

The precedence of operations can be overridden using parentheses. In particular, the precedence of juxtaposition implies that

\[
1/(2\pi),
\]

\[
1/(2\pi)(\pi+\pi) = (2\pi)^{-1}(\pi + \pi) \simeq 1,
\]

\[
\sin(2\pi) = \sin(2\pi) \neq 0,
\]

\[
2^{-2 \max(3,5)} = 2^2 \max(3,5) = 20,
\]

\[
1\text{in}/1\text{cm} = (1\text{in})/(1\text{cm}) = 2.54.
\]

Functions are called on the value of their argument, contrarily to \LaTeX macros.

9.3 Operations

We now present the various operations allowed in floating point expressions, from the lowest precedence to the highest. When used as a truth value, a floating point expression is \texttt{false} if it is \texttt{±0}, and \texttt{true} otherwise, including when it is \texttt{NaN} or a tuple such as (0, 0). Tuples are only supported to some extent by operations that work with truth values (?:, \texttt{11}, \texttt{&&}, \texttt{!}), by comparisons (\texttt{!\textless\textgreater\textlesseq?}), and by \texttt{+, -, *, /}. Unless otherwise specified, providing a tuple as an argument of any other operation yields the “invalid operation” exception and a \texttt{NaN} result.
The ternary operator `?:` results in `<operand 2>` if `<operand 1>` is true (not ±0), and `<operand 3>` if `<operand 1>` is false (±0). All three `<operands>` are evaluated in all cases; they may be tuples. The operator is right associative, hence

```
\fp_eval:n\{\langle\text{operand}_1\rangle \ ? \ \langle\text{operand}_2\rangle : \langle\text{operand}_3\rangle\}
```

first tests whether `1 + 3 > 4`; since this isn’t true, the branch following `:` is taken, and `2 + 4 > 5` is compared; since this is true, the branch before `:` is taken, and everything else is (evaluated then) ignored. That allows testing for various cases in a concise manner, with the drawback that all computations are made in all cases.

If `<operand 1>` is true (not ±0), use that value, otherwise the value of `<operand 2>`. Both `<operands>` are evaluated in all cases; they may be tuples. In `<operand 1> 11 <operand 2> 11 ... 11 <operand n>`, the first true (nonzero) `<operand>` is used and if all are zero the last one (±0) is used.

If `<operand 1>` is false (equal to ±0), use that value, otherwise the value of `<operand 2>`. Both `<operands>` are evaluated in all cases; they may be tuples. In `<operand 1> && <operand 2> && ... && <operand n>`, the first false (±0) `<operand>` is used and if none is zero the last one is used.

Each `<relation>` consists of a non-empty string of `<, =, >, and ?`, optionally preceded by !, and may not start with ?. This evaluates to +1 if all comparisons `<operand_i> <relation_i>` `<operand_{i+1}>` are true, and +0 otherwise. All `<operands>` are evaluated (once) in all cases. See \fp_compare:nTF for details.

Computes the sum or the difference of its two `<operands>`. The “invalid operation” exception occurs for ∞ − ∞. “Underflow” and “overflow” occur when appropriate. These operations supports the itemwise addition or subtraction of two tuples, but if they have a different number of items the “invalid operation” exception occurs and the result is NaN.
 Computes the product or the ratio of its two \(\textit{operands}\). The “invalid operation” exception occurs for \(\infty/\infty\), \(0/0\), or \(0 \times \infty\). “Division by zero” occurs when dividing a finite non-zero number by \(\pm 0\). “Underflow” and “overflow” occur when appropriate. When \(\textit{operand}_1\) is a tuple and \(\textit{operand}_2\) is a floating point number, each item of \(\textit{operand}_1\) is multiplied or divided by \(\textit{operand}_2\). Multiplication also supports the case where \(\textit{operand}_1\) is a floating point number and \(\textit{operand}_2\) a tuple. Other combinations yield an “invalid operation” exception and a NaN result.

 The unary + does nothing, the unary - changes the sign of the \(\textit{operand}\) (for a tuple, of all its components), and ! \(\textit{operand}\) evaluates to 1 if \(\textit{operand}\) is false (is \(\pm 0\)) and 0 otherwise (this is the not boolean function). Those operations never raise exceptions.

 Raises \(\textit{operand}_1\) to the power \(\textit{operand}_2\). This operation is right associative, hence \(2 \times 2 \times 3 = 256\). If \(\textit{operand}_1\) is negative or \(-0\) then: the result’s sign is + if the \(\textit{operand}_2\) is infinite and \((-1)^p\) if the \(\textit{operand}_2\) is \(p/5^q\) with \(p, q\) integers; the result is +0 if \(\text{abs}(\textit{operand}_1)^{\textit{operand}_2}\) evaluates to zero; in other cases the “invalid operation” exception occurs because the sign cannot be determined. “Division by zero” occurs when raising \(\pm 0\) to a finite strictly negative power. “Underflow” and “overflow” occur when appropriate. If either operand is a tuple, “invalid operation” occurs.

 Computes the absolute value of the \(\textit{fpexpr}\). If the operand is a tuple, “invalid operation” occurs. This operation does not raise exceptions in other cases. See also \(\textsf{abs(n)}\).

 Computes the exponential of the \(\textit{fpexpr}\). “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.

 Computes the factorial of the \(\textit{fpexpr}\). If the \(\textit{fpexpr}\) is an integer between \(-0\) and 3248 included, the result is finite and correctly rounded. Larger positive integers give \(+\infty\) with “overflow”, while \(\textsf{fact}(+\infty) = +\infty\) and \(\textsf{fact(nan)} = \text{nan}\) with no exception. All other inputs give NaN with the “invalid operation” exception.

 Computes the natural logarithm of the \(\textit{fpexpr}\). Negative numbers have no (real) logarithm, hence the “invalid operation” is raised in that case, including for \(\ln(-0)\). “Division by zero” occurs when evaluating \(\ln(+0) = -\infty\). “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.
\fp_eval:n { \logb( \fpexpr ) }  

Determines the exponent of the \langle \fpexpr \rangle, namely the floor of the base-10 logarithm of its absolute value. “Division by zero” occurs when evaluating \logb(\pm 0) = -\infty. Other special values are \logb(\pm \infty) = +\infty and \logb(\text{NaN}) = \text{NaN}. If the operand is a tuple or is \text{NaN}, then “invalid operation” occurs and the result is \text{NaN}.

\fp_eval:n { \max( \langle \fpexpr \rangle, \langle \fpexpr_2 \rangle, \ldots ) } \fp_eval:n { \min( \langle \fpexpr \rangle, \langle \fpexpr_2 \rangle, \ldots ) }  

Evaluates each \langle \fpexpr \rangle and computes the largest (smallest) of those. If any of the \langle \fpexpr \rangle is a \text{NaN} or tuple, the result is \text{NaN}. If any operand is a tuple, “invalid operation” occurs; these operations do not raise exceptions in other cases.

\fp_eval:n { \round( \langle \fpexpr \rangle ) } \fp_eval:n { \round( \langle \fpexpr_1 \rangle, \langle \fpexpr_2 \rangle ) } \fp_eval:n { \round( \langle \fpexpr_1 \rangle, \langle \fpexpr_2 \rangle, \langle \fpexpr_3 \rangle ) }  

Only \round accepts a third argument. Evaluates \langle \fpexpr_1 \rangle = x \text{ and } \langle \fpexpr_2 \rangle = n \text{ and } \langle \fpexpr_3 \rangle = t \text{ then rounds } x \text{ to } n \text{ places}. If \text{n} is an integer, this rounds \text{x} to a multiple of \text{10}^{-\text{n}}; if \text{n} = +\infty, this always yields \text{x}; if \text{n} = -\infty, this yields one of \pm 0, \pm \infty, \text{ or } \text{NaN}; if \text{n} = \text{NaN}, this yields \text{NaN}; if \text{n} is neither \pm \infty nor an integer, then an “invalid operation” exception is raised. When \langle \fpexpr_2 \rangle is omitted, \text{n} = 0, \text{i.e.}, \langle \fpexpr_1 \rangle is rounded to an integer. The rounding direction depends on the function.

- \round yields the multiple of \text{10}^{-\text{n}} closest to \text{x}, with ties (\text{x} half-way between two such multiples) rounded as follows. If \text{t} is \text{nan} (or not given) the even multiple is chosen (“ties to even”), if \text{t} = \pm 0 the multiple closest to 0 is chosen (“ties to zero”), if \text{t} is positive/negative the multiple closest to \infty/\text{-}\infty is chosen (“ties towards positive/negative infinity”).
- \floor yields the largest multiple of \text{10}^{-\text{n}} smaller or equal to \text{x} (“round towards negative infinity”);
- \ceil yields the smallest multiple of \text{10}^{-\text{n}} greater or equal to \text{x} (“round towards positive infinity”);
- \trunc yields a multiple of \text{10}^{-\text{n}} with the same sign as \text{x} and with the largest absolute value less than that of \text{x} (“round towards zero”).

“Overflow” occurs if \text{x} is finite and the result is infinite (this can only happen if \langle \fpexpr_2 \rangle < \text{-}9984). If any operand is a tuple, “invalid operation” occurs.

\fp_eval:n { \sign( \langle \fpexpr \rangle ) }  

Evaluates the \langle \fpexpr \rangle and determines its sign: +1 for positive numbers and for +\infty, -1 for negative numbers and for -\infty, \pm 0 for \pm 0, and \text{NaN} for \text{NaN}. If the operand is a tuple, “invalid operation” occurs. This operation does not raise exceptions in other cases.
\begin{verbatim}
\fp_eval:n { sin( \langle fpexpr \rangle ) }
\fp_eval:n { cos( \langle fpexpr \rangle ) }
\fp_eval:n { tan( \langle fpexpr \rangle ) }
\fp_eval:n { cot( \langle fpexpr \rangle ) }
\fp_eval:n { csc( \langle fpexpr \rangle ) }
\fp_eval:n { sec( \langle fpexpr \rangle ) }
\end{verbatim}

Computes the sine, cosine, tangent, cotangent, cosecant, or secant of the \langle fpexpr \rangle given in radians. For arguments given in degrees, see sind, cosd, etc. Note that since π is irrational, \( \sin(8\pi) \) is not quite zero, while its analogue \( \sin(8 \times 180) \) is exactly zero. The trigonometric functions are undefined for an argument of ±\( \infty \), leading to the “invalid operation” exception. Additionally, evaluating tangent, cotangent, cosecant, or secant at one of their poles leads to a “division by zero” exception. “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.

\begin{verbatim}
\fp_eval:n { sind( \langle fpexpr \rangle ) }
\fp_eval:n { cosd( \langle fpexpr \rangle ) }
\fp_eval:n { tand( \langle fpexpr \rangle ) }
\fp_eval:n { cotd( \langle fpexpr \rangle ) }
\fp_eval:n { cscd( \langle fpexpr \rangle ) }
\fp_eval:n { secd( \langle fpexpr \rangle ) }
\end{verbatim}

Computes the sine, cosine, tangent, cotangent, cosecant, or secant of the \langle fpexpr \rangle given in degrees. For arguments given in radians, see sin, cos, etc. Note that since π is irrational, \( \sin(8\pi) \) is not quite zero, while its analogue \( \sin(8 \times 180) \) is exactly zero. The trigonometric functions are undefined for an argument of ±\( \infty \), leading to the “invalid operation” exception. Additionally, evaluating tangent, cotangent, cosecant, or secant at one of their poles leads to a “division by zero” exception. “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.

\begin{verbatim}
\fp_eval:n { asin( \langle fpexpr \rangle ) }
\fp_eval:n { acos( \langle fpexpr \rangle ) }
\fp_eval:n { acsc( \langle fpexpr \rangle ) }
\fp_eval:n { asec( \langle fpexpr \rangle ) }
\end{verbatim}

Computes the arcsine, arccosine, arccosecant, or arcsecant of the \langle fpexpr \rangle and returns the result in radians, in the range \( [-\pi/2, \pi/2] \) for asin and acsc and \( [0, \pi] \) for acos and asec. For a result in degrees, use asind, etc. If the argument of asin or acos lies outside the range \( [-1, 1] \), or the argument of acsc or asec inside the range \( (-1, 1) \), an “invalid operation” exception is raised. “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.

\begin{verbatim}
\fp_eval:n { asind( \langle fpexpr \rangle ) }
\fp_eval:n { acosd( \langle fpexpr \rangle ) }
\fp_eval:n { acscd( \langle fpexpr \rangle ) }
\fp_eval:n { asecd( \langle fpexpr \rangle ) }
\end{verbatim}

Computes the arcsine, arccosine, arccosecant, or arcsecant of the \langle fpexpr \rangle and returns the result in degrees, in the range \( [-90, 90] \) for asin and acsc and \( [0, 180] \) for acos and asec. For a result in radians, use asin, etc. If the argument of asin or acos lies outside the range \( [-1, 1] \), or the argument of acsc or asec inside the range \( (-1, 1) \), an “invalid operation” exception is raised. “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.
\[ \text{atan} \]
\[ \text{acot} \]
\[ \text{sqrt} \]

Those functions yield an angle in radians: \text{atan} and \text{acot} are their analogs in degrees. The one-argument versions compute the arctangent or arccotangent of the \( \langle \text{fpexpr} \rangle \): arctangent takes values in the range \([−\pi/2, \pi/2]\), and arccotangent in the range \([0, \pi]\). The two-argument arctangent computes the angle in polar coordinates of the point with Cartesian coordinates \((\text{fpexpr}_2, \text{fpexpr}_1)\): this is the arctangent of \( \langle \text{fpexpr}_1 \rangle / \langle \text{fpexpr}_2 \rangle \), possibly shifted by \( \pi \) depending on the signs of \( \langle \text{fpexpr}_1 \rangle \) and \( \langle \text{fpexpr}_2 \rangle \). The two-argument arccotangent computes the angle in polar coordinates of the point \((\text{fpexpr}_1, \text{fpexpr}_2)\), equal to the arccotangent of \( \langle \text{fpexpr}_1 \rangle / \langle \text{fpexpr}_2 \rangle \), possibly shifted by \( \pi \). Both two-argument functions take values in the wider range \([−\pi, \pi]\). The ratio \( \langle \text{fpexpr}_1 \rangle / \langle \text{fpexpr}_2 \rangle \) need not be defined for the two-argument arctangent: when both expressions yield \pm0, or when both yield \pm\infty, the resulting angle is one of \( \{\pm\pi/4, \pm3\pi/4\} \) depending on signs. The “underflow” exception can occur. If any operand is a tuple, “invalid operation” occurs.

\[ \text{atan} \]
\[ \text{acot} \]
\[ \text{sqrt} \]

Computes the square root of the \( \langle \text{fpexpr} \rangle \). The “invalid operation” is raised when the \( \langle \text{fpexpr} \rangle \) is negative or is a tuple; no other exception can occur. Special values yield \( \sqrt{-0} = -0 \), \( \sqrt{+0} = +0 \), \( \sqrt{-\infty} = +\infty \) and \( \sqrt{\text{NaN}} = \text{NaN} \).
\fp_eval:n \{ \rand() \}

New: 2016-12-05

Produces a pseudo-random floating-point number (multiple of $10^{-16}$) between 0 included and 1 excluded. This is not available in older versions of \TeX. The random seed can be queried using \sys_rand_seed: and set using \sys_gset_rand_seed:n.

\TeXhackers note: This is based on pseudo-random numbers provided by the engine’s primitive \pduniformdeviate in pdf\TeX, \psuniformdeviate in up\TeX and \uniformdeviate in Lua\TeX and \X\TeX. The underlying code is based on Metapost, which follows an additive scheme recommended in Section 3.6 of “The Art of Computer Programming, Volume 2”.

While we are more careful than \uniformdeviate to preserve uniformity of the underlying stream of 28-bit pseudo-random integers, these pseudo-random numbers should of course not be relied upon for serious numerical computations nor cryptography.

\randint

\fp_eval:n \{ \randint{\langle fpexpr \rangle} \}
\fp_eval:n \{ \randint{\langle fpexpr \rangle_1, \langle fpexpr \rangle_2} \}

New: 2016-12-05

Produces a pseudo-random integer between 1 and $\langle fpexpr \rangle$ or between $\langle fpexpr \rangle_1$ and $\langle fpexpr \rangle_2$ inclusive. The bounds must be integers in the range $(-10^{16}, 10^{16})$ and the first must be smaller or equal to the second. See \rand for important comments on how these pseudo-random numbers are generated.

\inf The special values $+\infty$, $-\infty$, and NaN are represented as inf, -inf and nan (see \c_-inf_fp, \c_minus_inf_fp and \c_nan_fp).

\nan

\pi The value of \pi (see \c_pi_fp).

\deg The value of 1° in radians (see \c_one_degree_fp).
Those units of measurement are equal to their values in pt, namely

\begin{align*}
1 \text{in} &= 72.27 \text{pt} \\
1 \text{pt} &= 1 \text{pt} \\
1 \text{pc} &= 12 \text{pt} \\
1 \text{cm} &= \frac{1}{2.54} \text{in} = 28.45275590551181 \text{pt} \\
1 \text{mm} &= \frac{1}{25.4} \text{in} = 2.845275590551181 \text{pt} \\
1 \text{mm} &= 0.376065 \text{mm} = 1.07000856496063 \text{pt} \\
1 \text{cc} &= 12 \text{dd} = 12.8401027792756 \text{pt} \\
1 \text{nd} &= 0.375 \text{mm} = 1.066978346456693 \text{pt} \\
1 \text{cc} &= 12 \text{nd} = 12.80374015748031 \text{pt} \\
1 \text{bp} &= \frac{1}{72} \text{in} = 1.00375 \text{pt} \\
1 \text{sp} &= 2^{-16} \text{pt} = 1.52587890625c - 5 \text{pt}.
\end{align*}

The values of the (font-dependent) units \texttt{em} and \texttt{ex} are gathered from \TeX{} when the surrounding floating point expression is evaluated.

\begin{enumerate}
\item Other names for 1 and +0.
\end{enumerate}

\\begin{tabular}{ll}
\verb|\fp_abs:n| & \texttt{\{floating point expression\}} \\
& Evaluates the \texttt{\{floating point expression\}} as described for \texttt{\fp_eval:n} and leaves the \texttt{\{floating point expression\}} as described for \texttt{\fp_eval:n} and leaves the absolute value of the result in the input stream. If the argument is $\pm \infty$, NaN or a tuple, \texttt{\texttt{\fp_abs:n}} can be used; it accepts $\pm \infty$ and NaN as arguments.
\end{tabular}

\begin{tabular}{ll}
\verb|\fp_max:nn| & \texttt{\{fp expression 1\}} \{\texttt{fp expression 2\}} \\
\verb|\fp_min:nn| & \texttt{\{fp expression 1\}} \{\texttt{fp expression 2\}} \\
& Evaluates the \texttt{\{floating point expressions\}} as described for \texttt{\fp_eval:n} and leaves the resulting larger (\texttt{\texttt{\fp_max:nn}}) or smaller (\texttt{\fp_min:n}) value in the input stream. If the argument is a tuple, “invalid operation” occurs, but no other case raises exceptions. Within floating point expressions, \texttt{\texttt{\fp_max:nn}} and \texttt{\fp_min:n} can be used.
\end{tabular}

10 Disclaimer and roadmap

The package may break down if the escape character is among \texttt{0123456789_+}, or if it receives a \TeX{} primitive conditional affected by \texttt{\exp_not:N}.

The following need to be done. I’ll try to time-order the items.

\begin{itemize}
\item Function to count items in a tuple (and to determine if something is a tuple).
\item Decide what exponent range to consider.
\end{itemize}
• Support signalling \texttt{nan}.

• Modulo and remainder, and rounding function \texttt{quantize} (and its friends analogous to \texttt{trunc}, \texttt{ceil}, \texttt{floor}).

\begin{verbatim}
\texttt{\textbackslash fp\_format:nn \{fexpr\} \{format\}}, but what should \texttt{format} be? More general pretty printing?
\end{verbatim}

• Add \texttt{and}, \texttt{or}, \texttt{xor}? Perhaps under the names \texttt{all}, \texttt{any}, and \texttt{xor}?

• Add $\log(x, b)$ for logarithm of $x$ in base $b$.

• \texttt{hypot} (Euclidean length). Cartesian-to-polar transform.

• Hyperbolic functions \texttt{cosh, sinh, tanh}.

• Inverse hyperbolics.

• Base conversion, input such as \texttt{0xAB.CDEF}.

• Factorial (not with !), gamma function.

• Improve coefficients of the \texttt{sin} and \texttt{tan} series.

• Treat upper and lower case letters identically in identifiers, and ignore underscores.

• Add an \texttt{array(1,2,3)} and \texttt{i=complex(0,1)}.

• Provide an experimental \texttt{map} function? Perhaps easier to implement if it is a single character, \texttt{@sin(1,2)}?

• Provide an \texttt{isnan} function analogue of \texttt{\textbackslash fp\_if\_nan:nTF}?

• Support keyword arguments?

\texttt{Pgfmath} also provides box-measurements (depth, height, width), but boxes are not possible expandably.

Bugs, and tests to add.

• Check that functions are monotonic when they should.

• Add exceptions to ?:, !\texttt{<>?}, \&\&, ||, and !.

• Logarithms of numbers very close to 1 are inaccurate.

• When rounding towards $-\infty$, \texttt{\dim\_to\_fp:n \{0pt\}} should return $-0$, not $+0$.

• The result of $(\pm0) + (\pm0)$, of $x + (-x)$, and of $(\pm x) + x$ should depend on the rounding mode.

• 0e9999999999 gives a \TeX “number too large” error.

• Subnormals are not implemented.

Possible optimizations/improvements.

• Document that \texttt{l3trial/l3fp\_types} introduces tools for adding new types.

• In subsection 9.1, write a grammar.

217
• It would be nice if the parse auxiliaries for each operation were set up in the corresponding module, rather than centralizing in \texttt{l3fp-parse}.

• Some functions should get an \_o ending to indicate that they expand after their result.

• More care should be given to distinguish expandable/restricted expandable (auxiliary and internal) functions.

• The code for the ternary set of functions is ugly.

• There are many ~ missing in the doc to avoid bad line-breaks.

• The algorithm for computing the logarithm of the significand could be made to use a 5 terms Taylor series instead of 10 terms by taking $c = 2000/([200x]+1) \in [10, 95]$ instead of $c \in [1, 10]$. Also, it would then be possible to simplify the computation of $t$. However, we would then have to hard-code the logarithms of 44 small integers instead of 9.

• Improve notations in the explanations of the division algorithm (\texttt{l3fp-basics}).

• Understand and document \texttt{\_\_fp_basics_pack_weird_low:NNNNw} and \texttt{\_\_fp-_basics_pack_weird_high:NNNNNNNNw} better. Move the other \texttt{basics_pack} auxiliaries to \texttt{l3fp-aux} under a better name.

• Find out if underflow can really occur for trigonometric functions, and redoc as appropriate.

• Add bibliography. Some of Kahan’s articles, some previous \texttt{TeX} fp packages, the international standards,\ldots

• Also take into account the “inexact” exception?

• Support multi-character prefix operators (\textit{e.g.}, @/ or whatever)?
Part XXIV
The l3fparray package: fast global floating point arrays

1 l3fparray documentation

For applications requiring heavy use of floating points, this module provides arrays which can be accessed in constant time (contrast l3seq, where access time is linear). The interface is very close to that of l3intarray. The size of the array is fixed and must be given at point of initialisation

\fparray_new:Nn \fparray_new:Nn \fparray var \{ \text{size} \}

Evaluates the integer expression \text{(size)} and allocates an (floating point array variable) with that number of \text{(zero)} entries. The variable name should start with \text{\g_} because assignments are always global.

\fparray_count:N \fparray_count:N \fparray var

Expands to the number of entries in the (floating point array variable). This is performed in constant time.

\fparray_gset:Nnn \fparray_gset:Nnn \fparray var \{ \text{position} \} \{ \text{value} \}

Stores the result of evaluating the floating point expression \text{(value)} into the (floating point array variable) at the (integer expression) \text{(position)}. If the \text{(position)} is not between 1 and the \text{\fparray_count:N}, an error occurs. Assignments are always global.

\fparray_gzero:N \fparray_gzero:N \fparray var

Sets all entries of the (floating point array variable) to +0. Assignments are always global.

\fparray_item:Nn \fparray_item:Nn \fparray_item_to_tl:Nn \fparray_item_to_tl:Nn \fparray var \{ \text{position} \}

Applies \text{\fp_use:N} or \text{\fp_to_tl:N} (respectively) to the floating point entry stored at the (integer expression) \text{(position)} in the (floating point array variable). If the \text{(position)} is not between 1 and the \text{\fparray_count:N}, an error occurs.
Part XXV

The \texttt{l3sort} package

\section*{Sorting functions}

\subsection*{1 Controlling sorting}
\LaTeX{} comes with a facility to sort list variables (sequences, token lists, or comma-lists) according to some user-defined comparison. For instance,

\begin{verbatim}
\clist_set:Nn \l_foo_clist { 3 , 01 , -2 , 5 , +1 }
\clist_sort:Nn \l_foo_clist
{\int_compare:nNnTF { #1 } > { #2 }
{ \sort_return_swapped: }
{ \sort_return_same: }}
\end{verbatim}

results in \texttt{\l_foo_clist} holding the values \{\texttt{-2 , 01 , +1 , 3 , 5}\} sorted in non-decreasing order.

The code defining the comparison should call \texttt{\sort_return_swapped:} if the two items given as \texttt{#1} and \texttt{#2} are not in the correct order, and otherwise it should call \texttt{\sort_return_same:} to indicate that the order of this pair of items should not be changed.

For instance, a \langle\textit{comparison code}\rangle consisting only of \texttt{\sort_return_same:} with no test yields a trivial sort: the final order is identical to the original order. Conversely, using a \langle\textit{comparison code}\rangle consisting only of \texttt{\sort_return_swapped:} reverses the list (in a fairly inefficient way).

\texttt{\texttt{\LaTeX}hacker\texttt{note:}} The current implementation is limited to sorting approximately 20000 items (40000 in \texttt{\LaTeX}X), depending on what other packages are loaded.

Internally, the code from \texttt{l3sort} stores items in \texttt{\toks} registers allocated locally. Thus, the \langle\textit{comparison code}\rangle should not call \texttt{\newtoks} or other commands that allocate new \texttt{\toks} registers. On the other hand, altering the value of a previously allocated \texttt{\toks} register is not a problem.

\begin{verbatim}
\seq_sort:Nn \seq { ... \sort_return_same: or \sort_return_swapped: ... }
\end{verbatim}

 Indicates whether to keep the order or swap the order of two items that are compared in the sorting code. Only one of the \texttt{\sort_return_{\ldots}} functions should be used by the code, according to the results of some tests on the items \texttt{#1} and \texttt{#2} to be compared.
Part XXVI

The \texttt{l3tl-analysis} package: Analysing token lists

\section{\texttt{l3tl-analysis} documentation}

This module mostly provides internal functions for use in the \texttt{l3regex} module. However, it provides as a side-effect a user debugging function, very similar to the \texttt{ShowTokens} macro from the \texttt{ted} package.

\begin{verbatim}
\tl_analysis_show:n \{\token\}
\tl_analysis_show:N \tl_analysis_show:n
\end{verbatim}

Displays to the terminal the detailed decomposition of the \texttt{\token} into tokens, showing the category code of each character token, the meaning of control sequences and active characters, and the value of registers.

\begin{verbatim}
\tl_analysis_map_inline:nn \{\token\} \{\inline\}
\tl_analysis_map_inline:Nn \tl_analysis_map_inline:nn
\end{verbatim}

New: 2018-04-09

Applies the \texttt{\inline} to each individual \texttt{\token} in the \texttt{\token}. The \texttt{\inline} receives three arguments:

- \texttt{\tokens}, which both \texttt{o-expand} and \texttt{x-expand} to the \texttt{\token}. The detailed form of \texttt{\token} may change in later releases.

- \texttt{\char code}, a decimal representation of the character code of the token, \texttt{\texttt{\char code}} if it is a control sequence (with \texttt{\catcode} 0).

- \texttt{\catcode}, a capital hexadecimal digit which denotes the category code of the \texttt{\token} (0: control sequence, 1: begin-group, 2: end-group, 3: math shift, 4: alignment tab, 6: parameter, 7: superscript, 8: subscript, A: space, B: letter, C:other, D:active).

As all other mappings the mapping is done at the current group level, \textit{i.e.} any local assignments made by the \texttt{\inline} remain in effect after the loop.
Part XXVII

The l3regex package: Regular expressions in \TeX

The l3regex package provides regular expression testing, extraction of submatches, splitting, and replacement, all acting on token lists. The syntax of regular expressions is mostly a subset of the pcre syntax (and very close to POSIX), with some additions due to the fact that \TeX manipulates tokens rather than characters. For performance reasons, only a limited set of features are implemented. Notably, back-references are not supported.

Let us give a few examples. After
\begin{verbatim}
\tl_set:Nn \l_my_tl { That~cat. }
\regex_replace_once:nnN { at } { is } \l_my_tl
\end{verbatim}
the token list variable \emph{\l_my_tl} holds the text “This \textit{cat.}.”, where the first occurrence of “\textit{at}” was replaced by “\textit{is}”. A more complicated example is a pattern to emphasize each word and add a comma after it:
\begin{verbatim}
\regex_replace_all:nnN { \w+ } { \c{emph}\cB\{ \0 \cE\} , } \l_my_tl
\end{verbatim}
The \emph{\w} sequence represents any “word” character, and + indicates that the \emph{\w} sequence should be repeated as many times as possible (at least once), hence matching a word in the input token list. In the replacement text, \emph{\0} denotes the full match (here, a word). The command \emph{\c{emph}} is inserted using \emph{\c{emph}}, and its argument \emph{\0} is put between \emph{\cB\{ and \cE\}}.

If a regular expression is to be used several times, it can be compiled once, and stored in a regex variable using \regex_const:Nn. For example,
\begin{verbatim}
\regex_const:Nn \c_foo_regex { \c{begin} \cB. \c[^BE].* \cE. }
\end{verbatim}
stores in \emph{\c_foo_regex} a regular expression which matches the starting marker for an environment: \textbf{\c{begin}}, followed by a begin-group token (\emph{\cB.}), then any number of tokens which are neither begin-group nor end-group character tokens (\textbf{\c[^BE].*}), ending with an end-group token (\emph{\cE.}). As explained in the next section, the parentheses “capture” the result of \emph{\c[^BE].*}, giving us access to the name of the environment when doing replacements.

1 Syntax of regular expressions

We start with a few examples, and encourage the reader to apply \regex_show:n to these regular expressions.

- Cat matches the word “Cat” capitalized in this way, but also matches the beginning of the word “Cattle”: use \texttt{\bCat\b} to match a complete word only.

- [abc] matches one letter among “a”, “b”, “c”: the pattern (a|b|c) matches the same three possible letters (but see the discussion of submatches below).

- [A-Za-z]* matches any number (due to the quantifier *) of Latin letters (not accented).
• \c{[A-Za-z]*} matches a control sequence made of Latin letters.
• \_\{-\_\}*\_ matches an underscore, any number of characters other than underscore, and another underscore; it is equivalent to \_\.*?\_ where . matches arbitrary characters and the lazy quantifier *? means to match as few characters as possible, thus avoiding matching underscores.
• \[\+-\]d+ matches an explicit integer with at most one sign.
• \[\+-\]d+\_\* matches an explicit integer with any number of + and – signs, with spaces allowed except within the mantissa, and surrounded by spaces.
• \[\+-\]d*\.d+\_\* matches an explicit integer or decimal number; using \[.,]\ instead of \. would allow the comma as a decimal marker.
• \[\+-\]d*\.d+\_\*((?i)pt|in|[cm]m|ex|\[bs\]p|\[dn\]d|\[pcn\]c)\_\* matches an explicit dimension with any unit that \TeX{} knows, where (?i) means to treat lowercase and uppercase letters identically.
• \[\+-\]d*\((?i)nan|inf|\(\d+|\d*\.\d+)\)\_\* matches an explicit floating point number or the special values \texttt{nan} and \texttt{inf} (with signs and spaces allowed).
• \[\+-\]d*\cC.\_\* matches an explicit integer or control sequence (without checking whether it is an integer variable).
• \G.*?\K at the beginning of a regular expression matches and discards (due to \G) everything between the end of the previous match (\G) and what is matched by the rest of the regular expression; this is useful in \texttt{\regex_replace_all:nnN} when the goal is to extract matches or submatches in a finer way than with \texttt{\regex_extract_all:nnN}.

While it is impossible for a regular expression to match only integer expressions, \[\+-\]d*\(\(\d+|\d*\.\d+)\)\_\* matches among other things all valid integer expressions (made only with explicit integers). One should follow it with further testing.

Most characters match exactly themselves, with an arbitrary category code. Some characters are special and must be escaped with a backslash (e.g., \* matches a star character). Some escape sequences of the form backslash–letter also have a special meaning (for instance \texttt{\lodash} matches any digit). As a rule,

• every alphanumeric character (A–Z, a–z, 0–9) matches exactly itself, and should not be escaped, because \A, \B, . . . have special meanings;
• non-alphanumeric printable ascii characters can (and should) always be escaped: many of them have special meanings (e.g., use \texttt{\c{,}}, \texttt{\c{?}}, \texttt{\c{.}});
• spaces should always be escaped (even in character classes);
• any other character may be escaped or not, without any effect: both versions match exactly that character.
Note that these rules play nicely with the fact that many non-alphanumeric characters are difficult to input into \TeX under normal category codes. For instance, \abc\% matches the characters \abc (with arbitrary category codes), but does not match the control sequence \abc followed by a percent character. Matching control sequences can be done using the \c{(regex)} syntax (see below).

Any special character which appears at a place where its special behaviour cannot apply matches itself instead (for instance, a quantifier appearing at the beginning of a string), after raising a warning.

Characters.

\x{hh...} Character with hex code hh...

\xhh Character with hex code hh.

\a Alarm (hex 07).
\e Escape (hex 1B).
\f Form-feed (hex 0C).
\n New line (hex 0A).
\r Carriage return (hex 0D).
\t Horizontal tab (hex 09).

Character types.

• A single period matches any token.
\d Any decimal digit.
\h Any horizontal space character, equivalent to [\ ?]: space and tab.
\s Any space character, equivalent to [\ ?J\ ?L\ ?M].
\v Any vertical space character, equivalent to [\ ?J\ ?K\ ?L\ ?M]. Note that \?K is a vertical space, but not a space, for compatibility with Perl.
\w Any word character, i.e., alphanumerics and underscore, equivalent to the explicit class [A-Za-z0-9\_].
\D Any token not matched by \d.
\H Any token not matched by \h.
\N Any token other than the \n character (hex 0A).
\S Any token not matched by \s.
\V Any token not matched by \v.
\W Any token not matched by \w.

Of those, .\, \D, \H, \N, \S, \V, and \W match arbitrary control sequences. Character classes match exactly one token in the subject.

[. . . ] Positive character class. Matches any of the specified tokens.
[^...] Negative character class. Matches any token other than the specified characters.

x-y Within a character class, this denotes a range (can be used with escaped characters).

[:\{name\}:] Within a character class (one more set of brackets), this denotes the POSIX character class \{name\}, which can be alnum, alpha, ascii, blank, cntrl, digit, graph, lower, print, punct, space, upper, word, or xdigit.

[:~\{name\}:] Negative POSIX character class.

For instance, [a-oq-z\cC.] matches any lowercase latin letter except p, as well as control sequences (see below for a description of \c).

Quantifiers (repetition).

? 0 or 1, greedy.

?? 0 or 1, lazy.

* 0 or more, greedy.

*? 0 or more, lazy.

+ 1 or more, greedy.

+? 1 or more, lazy.

\{n\} Exactly n.

\{n,\} n or more, greedy.

\{n,\}? n or more, lazy.

\{n,m\} At least n, no more than m, greedy.

\{n,m\}? At least n, no more than m, lazy.

Anchors and simple assertions.

\b Word boundary: either the previous token is matched by \w and the next by \W, or the opposite. For this purpose, the ends of the token list are considered as \W.

\B Not a word boundary: between two \w tokens or two \W tokens (including the boundary).

\^ or \A Start of the subject token list.

\$ \Z or \z End of the subject token list.

\G Start of the current match. This is only different from ^ in the case of multiple matches: for instance \[regex_count:nnN \{ \G a \} \{ aaba \} \l_tmpa_int\] yields 2, but replacing \G by ^ would result in \l_tmpa_int holding the value 1.

Alternation and capturing groups.

A|B|C Either one of A, B, or C.

(....) Capturing group.

(?:....) Non-capturing group.
(?!...) Non-capturing group which resets the group number for capturing groups in each alternative. The following group is numbered with the first unused group number.

The \c escape sequence allows to test the category code of tokens, and match control sequences. Each character category is represented by a single uppercase letter:

- C for control sequences;
- B for begin-group tokens;
- E for end-group tokens;
- M for math shift;
- T for alignment tab tokens;
- P for macro parameter tokens;
- U for superscript tokens (up);
- D for subscript tokens (down);
- S for spaces;
- L for letters;
- O for others; and
- A for active characters.

The \c escape sequence is used as follows.

\c\{\textit{regex}\} A control sequence whose csname matches the \(\textit{regex}\), anchored at the beginning and end, so that \c\{\textit{begin}\} matches exactly \textit{begin} and nothing else.

\cX Applies to the next object, which can be a character, character property, class, or group, and forces this object to only match tokens with category \texttt{X} (any of \texttt{CBEMTPUDSLOA}). For instance, \c\{L[A-Z]\d\} matches uppercase letters and digits of category code letter, \c\{C\}. matches any control sequence, and \c\{O(abc)\} matches \texttt{abc} where each character has category other.

\c\{XYZ\} Applies to the next object, and forces it to only match tokens with category \texttt{X}, \texttt{Y}, or \texttt{Z} (each being any of \texttt{CBEMTPUDSLOA}). For instance, \c\{LSO\}(..) matches two tokens of category letter, space, or other.

\c\{^XYZ\} Applies to the next object and prevents it from matching any token with category \texttt{X}, \texttt{Y}, or \texttt{Z} (each being any of \texttt{CBEMTPUDSLOA}). For instance, \c\{^O\}\d matches digits which have any category different from other.

The category code tests can be used inside classes; for instance, \c\{c[0-9]\c\{[L0-9][A-F]\}\}\ matches what \TeX\ considers as hexadecimal digits, namely digits with category other, or uppercase letters from A to F with category either letter or other. Within a group affected by a category code test, the outer test can be overridden by a nested test: for instance, \c\{L(ab)c0\c\{c\cd\} matches \texttt{abcd} where all characters are of category letter, except \* which has category other.

The \u escape sequence allows to insert the contents of a token list directly into a regular expression or a replacement, avoiding the need to escape special characters.
Namely, `\u{tl var name}` matches the exact contents of the token list `(tl var)`. Within a `\c{...}` control sequence matching, the `\u` escape sequence only expands its argument once, in effect performing `\tl_to_str:v`. Quantifiers are not supported directly: use a group.

The option `(?i)` makes the match case insensitive (identifying `A–Z` with `a–z`; no Unicode support yet). This applies until the end of the group in which it appears, and can be reverted using `(?-i)`. For instance, in `(?i)(a(?-i)b|c)d`, the letters `a` and `d` are affected by the `i` option. Characters within ranges and classes are affected individually: `(?i)[Y-\]` is equivalent to `[YZ][\yz]`, and `(?i)[^aeiou]` matches any character which is not a vowel. Neither character properties, nor `\c{...}` nor `\u{...}` are affected by the `i` option.

In character classes, only `[`, `^`, `\`, and spaces are special, and should be escaped. Other non-alphanumeric characters can still be escaped without harm. Any escape sequence which matches a single character (`\d`, `\D`, etc.) is supported in character classes. If the first character is `^`, then the meaning of the character class is inverted; `^` appearing anywhere else in the range is not special. If the first character (possibly following a leading `^`) is `]` then it does not need to be escaped since ending the range there would make it empty. Ranges of characters can be expressed using `^`, for instance, `[\D 0-5]` and `[^6-9]` are equivalent.

Capturing groups are a means of extracting information about the match. Parenthesized groups are labelled in the order of their opening parenthesis, starting at 1. The contents of those groups corresponding to the “best” match (leftmost longest) can be extracted and stored in a sequence of token lists using for instance `\regex_extract_once:nnNTF`. The `\K` escape sequence resets the beginning of the match to the current position in the token list. This only affects what is reported as the full match. For instance,

```
\regex_extract_all:nnN { a \K . } { a123aaxyz } \l_foo_seq
```

results in `\l_foo_seq` containing the items `{1}` and `{a}`: the true matches are `{a1}` and `{aa}`, but they are trimmed by the use of `\K`. The `\K` command does not affect capturing groups: for instance,

```
\regex_extract_once:nnM { (. \K c)+ \d } { acbc3 } \l_foo_seq
```

results in `\l_foo_seq` containing the items `{c3}` and `{bc}`: the true match is `{acbc3}`, with first submatch `{bc}`, but `\K` resets the beginning of the match to the last position where it appears.

### 2 Syntax of the replacement text

Most of the features described in regular expressions do not make sense within the replacement text. Backslash introduces various special constructions, described further below:

- `\0` is the whole match;
- `\1` is the submatch that was matched by the first (capturing) group `(....)`; similarly for `\2`, `. . .`, `\9` and `\g{number}`;
- `\ ` inserts a space (spaces are ignored when not escaped);
• \a, \e, \f, \n, \r, \t, \xhh, \x{hhh} correspond to single characters as in regular expressions;
• \c\{cs name\} inserts a control sequence;
• \c\{category\}\{character\} (see below);
• \u\{\{tl var name\}\} inserts the contents of the \{tl var\} (see below).

Characters other than backslash and space are simply inserted in the result (but since the replacement text is first converted to a string, one should also escape characters that are special for \TeX, for instance use \#). Non-alphanumeric characters can always be safely escaped with a backslash.

For instance,

\tl_set:Nn \l_my_tl { Hello,-world! }
\regex_replace_all:nnN { ([er]?l|o) . } { (\0--\1) } \l_my_tl

results in \l_my_tl holding H(ell--el)(o,--o) w(or--o)(l--l)!

The submatches are numbered according to the order in which the opening parenthesis of capturing groups appear in the regular expression to match. The \text{n}-th submatch is empty if there are fewer than \text{n} capturing groups or for capturing groups that appear in alternatives that were not used for the match. In case a capturing group matches several times during a match (due to quantifiers) only the last match is used in the replacement text. Submatches always keep the same category codes as in the original token list.

The characters inserted by the replacement have category code 12 (other) by default, with the exception of space characters. Spaces inserted through \u have category code 10, while spaces inserted through \x20 or \x{20} have category code 12. The escape sequence \c allows to insert characters with arbitrary category codes, as well as control sequences.

\cX(...) Produces the characters “...” with category X, which must be one of CBEMTPUDSLOA as in regular expressions. Parentheses are optional for a single character (which can be an escape sequence). When nested, the innermost category code applies, for instance \cL(Hello\cS\ world)! gives this text with standard category codes.

\c\{\text\} Produces the control sequence with csname \text. The \text may contain references to the submatches \0, \1, and so on, as in the example for \u below.

The escape sequence \u\{\{tl var name\}\} allows to insert the contents of the token list with name \{tl var name\} directly into the replacement, giving an easier control of category codes. When nested in \c\{\text\}\} and \u\{\text\}\} constructions, the \u and \c escape sequences perform \tl_to_str:v, namely extract the value of the control sequence and turn it into a string. Matches can also be used within the arguments of \c and \u. For instance,

\tl_set:Nn \l_my_one_tl { first }
\tl_set:Nn \l_my_two_tl { \emph{second} }\}
\tl_set:Nn \l_my_tl { one , two , one , one }
\regex_replace_all:nnN { [-,]+ } { \u\{l_my\_0_tl\} } \l_my_tl

results in \l_my_tl holding first,\emph{second},first,first.
3 Pre-compiling regular expressions

If a regular expression is to be used several times, it is better to compile it once rather than doing it each time the regular expression is used. The compiled regular expression is stored in a variable. All of the \texttt{\texttt{l3regex}} module’s functions can be given their regular expression argument either as an explicit string or as a compiled regular expression.

\begin{verbatim}
\regex_new:N ⟨regex var⟩
\end{verbatim}

Creates a new \texttt{⟨regex var⟩} or raises an error if the name is already taken. The declaration is global. The \texttt{⟨regex var⟩} is initially such that it never matches.

\begin{verbatim}
\regex_set:Nn ⟨regex var⟩ {⟨regex⟩}
\end{verbatim}

Stores a compiled version of the \texttt{⟨regular expression⟩} in the \texttt{⟨regex var⟩}. For instance, this function can be used as

\begin{verbatim}
\regex_new:N \l_my_regex
\regex_set:Nn \l_my_regex { my\ (simple\ )? reg(ex|ular\ expression) }
\end{verbatim}

The assignment is local for \texttt{\regex_set:Nn} and global for \texttt{\regex_gset:Nn}. Use \texttt{\regex_const:Nn} for compiled expressions which never change.

\begin{verbatim}
\regex_show:n {⟨regex⟩}
\end{verbatim}

Shows how \texttt{l3regex} interprets the \texttt{⟨regex⟩}. For instance, \texttt{\regex_show:n {\A X|Y}} shows

\begin{itemize}
  \item \texttt{+-branch}
  \item anchor at start (\texttt{\A})
  \item char code 88
  \item \texttt{+-branch}
  \item char code 89
\end{itemize}

indicating that the anchor \texttt{\A} only applies to the first branch: the second branch is not anchored to the beginning of the match.

4 Matching

All regular expression functions are available in both \texttt{:n} and \texttt{:N} variants. The former require a “standard” regular expression, while the later require a compiled expression as generated by \texttt{\regex_(g)set:Nn}.

\begin{verbatim}
\regex_match:nnTF {⟨regex⟩} {⟨token list⟩} {⟨true code⟩} {⟨false code⟩}
\end{verbatim}

Tests whether the \texttt{⟨regular expression⟩} matches any part of the \texttt{⟨token list⟩}. For instance,

\begin{verbatim}
\regex_match:nnTF { b [cde]* } { abedcx } { TRUE } { FALSE }
\regex_match:nnTF { [b-dq-w] } { example } { TRUE } { FALSE }
\end{verbatim}

leaves TRUE then FALSE in the input stream.
The \texttt{\regex_count:nnN} macro sets an integer variable equal to the number of times a regular expression appears in a token list. The search starts by finding the left-most longest match, respecting greedy and lazy (non-greedy) operators. Then the search starts again from the character following the last character of the previous match, until reaching the end of the token list. Infinite loops are prevented in the case where the regular expression can match an empty token list: then we count one match between each pair of characters.

For instance,

\begin{verbatim}
\int_new:N \l_foo_int
\regex_count:nnN { (b+|c) } { abbababcbb } \l_foo_int
\end{verbatim}

results in \texttt{\l_foo_int} taking the value 5.

\section{Submatch extraction}

\texttt{\regex_extract_once:nnN} finds the first match of the regular expression in the token list. If it exists, the match is stored as the first item of the sequence, and further items are the contents of capturing groups, in the order of their opening parenthesis. The sequence is assigned locally. If there is no match, the sequence is cleared. The testing versions insert the \texttt{true code} into the input stream if a match was found, and the \texttt{false code} otherwise.

For instance, assume that you type

\begin{verbatim}
\regex_extract_once:nnNTF { \A(La)?TeX(!*)\Z } { LaTeX!!! } \l_foo_seq
\end{verbatim}

Then the regular expression (anchored at the start with \texttt{\A} and at the end with \texttt{\Z}) must match the whole token list. The first capturing group, \texttt{(La)?}, matches \texttt{La}, and the second capturing group, \texttt{(!*)}, matches \texttt{!!!}. Thus, \texttt{\l_foo_seq} contains as a result the items \texttt{LaTeX!!!}, \texttt{La}, and \texttt{!!!}, and the \texttt{true} branch is left in the input stream. Note that the \texttt{n}-th item of \texttt{\l_foo_seq}, as obtained using \texttt{\seq_item:Nn}, correspond to the submatch numbered \texttt{(n-1)} in functions such as \texttt{\regex_replace_once:nnN}.

\texttt{\regex_extract_all:nnNTF} finds all matches of the regular expression in the token list, and stores all the submatch information in a single sequence (concatenating the results of multiple \texttt{\regex_extract_once:nnN} calls). The \texttt{\seq_var} is assigned locally. If there is no match, the \texttt{\seq_var} is cleared. The testing versions insert the \texttt{true code} into the input stream if a match was found, and the \texttt{false code} otherwise. For instance, assume that you type

\begin{verbatim}
\regex_extract_all:nnNTF { \w+ } { Hello,-world! } \l_foo_seq
\end{verbatim}

Then the regular expression matches twice, the resulting sequence contains the two items \texttt{Hello} and \texttt{world}, and the \texttt{true} branch is left in the input stream.
Splits the ⟨token list⟩ into a sequence of parts, delimited by matches of the ⟨regular expression⟩. If the ⟨regular expression⟩ has capturing groups, then the token lists that they match are stored as items of the sequence as well. The assignment to ⟨seq var⟩ is local. If no match is found the resulting ⟨seq var⟩ has the ⟨token list⟩ as its sole item. If the ⟨regular expression⟩ matches the empty token list, then the ⟨token list⟩ is split into single tokens. The testing versions insert the ⟨true code⟩ into the input stream if a match was found, and the ⟨false code⟩ otherwise. For example, after

\seq_new:N \l_path_seq
\regex_split:nnNTF { / } { the/path/for/this/file.tex } \l_path_seq
{ true } { false }

the sequence \l_path_seq contains the items {the}, {path}, {for}, {this}, and {file.tex}, and the true branch is left in the input stream.

### 6 Replacement

Searches for the ⟨regular expression⟩ in the ⟨token list⟩ and replaces the first match with the ⟨replacement⟩. The result is assigned locally to ⟨tl var⟩. In the ⟨replacement⟩, \0 represents the full match, \1 represent the contents of the first capturing group, \2 of the second, etc.

\regex_replace_all:nnNTF { / } { \the/path/for/this/file.tex } \l_path_seq
{ true } { false }

Replaces all occurrences of the ⟨regular expression⟩ in the ⟨token list⟩ by the ⟨replacement⟩, where \0 represents the full match, \1 represent the contents of the first capturing group, \2 of the second, etc. Every match is treated independently, and matches cannot overlap. The result is assigned locally to ⟨tl var⟩.

### 7 Constants and variables

\l_tmpa_regex
\l_tmpb_regex

Scratch regex for local assignment. These are never used by the kernel code, and so are safe for use with any \texttt{\LaTeX}3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\g_tmpa_regex
\g_tmpb_regex

Scratch regex for global assignment. These are never used by the kernel code, and so are safe for use with any \texttt{\LaTeX}3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
8 Bugs, misfeatures, future work, and other possibilities

The following need to be done now.

- Rewrite the documentation in a more ordered way, perhaps add a BNF?
  Additional error-checking to come.
- Clean up the use of messages.
- Cleaner error reporting in the replacement phase.
- Add tracing information.
- Detect attempts to use back-references and other non-implemented syntax.
- Test for the maximum register $\texttt{\_c\_max\_register\_int}$.
- Find out whether the fact that $\texttt{\W}$ and friends match the end-marker leads to bugs.
  Possibly update $\texttt{\_\_regex\_item\_reverse:n}$.
- The empty cs should be matched by $\texttt{\c\{\}}$, not by $\texttt{\c\{csname\?endcsname\s?\}}$.
  Code improvements to come.
- Shift arrays so that the useful information starts at position 1.
- Only build $\texttt{\c\{\ldots\}}$ once.
- Use arrays for the left and right state stacks when compiling a regex.
- Should $\texttt{\_\_regex\_action\_free\_group:n}$ only be used for greedy $\{n,\}$ quantifier?
  (I think not.)
- Quantifiers for $\texttt{\u}$ and assertions.
- When matching, keep track of an explicit stack of $\texttt{\_current\_state}$ and $\texttt{\_current\_-submatches}$.
- If possible, when a state is reused by the same thread, kill other subthreads.
- Use an array rather than $\texttt{\l\_\_regex\_balance\_tl}$ to build the function $\texttt{\_\_regex\_replacement\_balance\_one\_match:n}$.
- Reduce the number of epsilon-transitions in alternatives.
- Optimize simple strings: use less states ($\texttt{abcade}$ should give two states, for $\texttt{abc}$ and $\texttt{ade}$). [Does that really make sense?]
- Optimize groups with no alternative.
- Optimize states with a single $\texttt{\_\_regex\_action\_free:n}$.
- Optimize the use of $\texttt{\_\_regex\_action\_success}$: by inserting it in state 2 directly instead of having an extra transition.
- Optimize the use of $\texttt{\int\_step\ldots}$ functions.
• Groups don’t capture within regexes for csnames; optimize and document.
• Better “show” for anchors, properties, and catcode tests.
• Does \K really need a new state for itself?
• When compiling, use a boolean in_cs and less magic numbers.
• Instead of checking whether the character is special or alphanumeric using its character code, check if it is special in regexes with \cs_if_exist tests.

The following features are likely to be implemented at some point in the future.

• General look-ahead/behind assertions.
• Regex matching on external files.
• Conditional subpatterns with look ahead/behind: “if what follows is […], then […].”
• (*..) and (?..) sequences to set some options.
• UTF-8 mode for pdfTeX.
• Newline conventions are not done. In particular, we should have an option for . not to match newlines. Also, \A should differ from ^, and \Z, \z and $ should differ.
• Unicode properties: \p{..} and \P{..}; \X which should match any “extended” Unicode sequence. This requires to manipulate a lot of data, probably using tree-boxes.
• Provide a syntax such as \ur{l_my_regex} to use an already-compiled regex in a more complicated regex. This makes regexes more easily composable.
• Allowing \u{l_my_tl} in more places, for instance as the number of repetitions in a quantifier.

The following features of pcre or Perl may or may not be implemented.

• Callout with (?C…) or other syntax: some internal code changes make that possible, and it can be useful for instance in the replacement code to stop a regex replacement when some marker has been found; this raises the question of a potential \regex_break: and then of playing well with \tl_map_break: called from within the code in a regex. It also raises the question of nested calls to the regex machinery, which is a problem since \fontdimen are global.
• Conditional subpatterns (other than with a look-ahead or look-behind condition): this is non-regular, isn’t it?
• Named subpatterns: \TeX programmers have lived so far without any need for named macro parameters.

The following features of pcre or Perl will definitely not be implemented.

• Back-references: non-regular feature, this requires backtracking, which is prohibitively slow.
• Recursion: this is a non-regular feature.

• Atomic grouping, possessive quantifiers: those tools, mostly meant to fix catastrophic backtracking, are unnecessary in a non-backtracking algorithm, and difficult to implement.

• Subroutine calls: this syntactic sugar is difficult to include in a non-backtracking algorithm, in particular because the corresponding group should be treated as atomic.

• Backtracking control verbs: intrinsically tied to backtracking.

• \ddd, matching the character with octal code ddd: we already have \x{...} and the syntax is confusingly close to what we could have used for backreferences (\1, \2, ...), making it harder to produce useful error message.

• \cx, similar to \TeX’s own \^^x.

• Comments: \TeX already has its own system for comments.

• \Q...\E escaping: this would require to read the argument verbatim, which is not in the scope of this module.

• \C single byte in UTF-8 mode: \TeX and \LaTeX serve us characters directly, and splitting those into bytes is tricky, encoding dependent, and most likely not useful anyways.
Part XXVIII

The \texttt{l3box} package

Boxes

There are three kinds of box operations: horizontal mode denoted with prefix \texttt{\hbox}_{},
vertical mode with prefix \texttt{\vbox}_{}, and the generic operations working in both modes with
prefix \texttt{\box}_{}.

1 Creating and initialising boxes

\texttt{\box\_new:N} \langle \textit{box} \rangle

Creates a new \langle \textit{box} \rangle or raises an error if the name is already taken. The declaration is
global. The \langle \textit{box} \rangle is initially void.

\texttt{\box\_clear:N} \langle \textit{box} \rangle

Clears the content of the \langle \textit{box} \rangle by setting the box equal to \texttt{\c\_empty\_box}.

\texttt{\box\_clear\_new:N} \langle \textit{box} \rangle

Ensures that the \langle \textit{box} \rangle exists globally by applying \texttt{\box\_new:N} if necessary, then applies
\texttt{\box\_\(g\)\_clear:N} to leave the \langle \textit{box} \rangle empty.

\texttt{\box\_set\_eq:NN} \langle \textit{box}1 \rangle \langle \textit{box}2 \rangle

Sets the content of \langle \textit{box}1 \rangle equal to that of \langle \textit{box}2 \rangle.

\texttt{\box\_if\_exist_p:N * \langle \textit{box} \rangle}
\texttt{\box\_if\_exist_p:c * \langle \textit{box} \rangle}
\texttt{\box\_if\_exist:NTF * \langle \textit{box} \rangle}
\texttt{\box\_if\_exist:cfT *}

Tests whether the \langle \textit{box} \rangle is currently defined. This does not check that the \langle \textit{box} \rangle really is
a box.

2 Using boxes

\texttt{\box\_use:N} \langle \textit{box} \rangle

Inserts the current content of the \langle \textit{box} \rangle onto the current list for typesetting. An error is
raised if the variable does not exist or if it is invalid.

\texttt{\TeX{h}ackers note:} This is the \TeX{} primitive \texttt{\copy{}}.

\new: 2012-03-03
\box_move_right:nn \box_move_left:nn

This function operates in vertical mode, and inserts the material specified by the \textit{box function} such that its reference point is displaced horizontally by the given \textit{dimexpr} from the reference point for typesetting, to the right or left as appropriate. The \textit{box function} should be a box operation such as \texttt{\box_use:N \box} or a “raw” box specification such as \texttt{\vbox:n \{ xyz \}}.

\box_move_up:nn \box_move_down:nn

This function operates in horizontal mode, and inserts the material specified by the \textit{box function} such that its reference point is displaced vertically by the given \textit{dimexpr} from the reference point for typesetting, up or down as appropriate. The \textit{box function} should be a box operation such as \texttt{\box_use:N \box} or a “raw” box specification such as \texttt{\vbox:n \{ xyz \}}.

3 Measuring and setting box dimensions

\box_dp:N \box_dp:c

Calculates the depth (below the baseline) of the \textit{box} in a form suitable for use in a \textit{dimension expression}.

\TeXhacksnote This is the \TeX primitive \texttt{\dp}.

\box_ht:N \box_ht:c

Calculates the height (above the baseline) of the \textit{box} in a form suitable for use in a \textit{dimension expression}.

\TeXhacksnote This is the \TeX primitive \texttt{\ht}.

\box_wd:N \box_wd:c

Calculates the width of the \textit{box} in a form suitable for use in a \textit{dimension expression}.

\TeXhacksnote This is the \TeX primitive \texttt{\wd}.

\box_set_dp:Nn \box_set_dp:cn \box_gset_dp:Nn \box_gset_dp:cn

Set the depth (below the baseline) of the \textit{box} to the value of the \textit{dimension expression}.

Updated: 2019-01-22

\box_set_ht:Nn \box_set_ht:cn \box_gset_ht:Nn \box_gset_ht:cn

Set the height (above the baseline) of the \textit{box} to the value of the \textit{dimension expression}.

Updated: 2019-01-22
\box_set_wd:Nn \box_set_wd:cn \box_gset_wd:Nn \box_gset_wd:cn

Updated: 2019-01-22

4 Box conditionals

\box_if_empty_p:N \box_if_empty_p:c \box_if_empty:NTF \box_if_empty:TF

Tests if \texttt{box} is a empty (equal to \texttt{c_empty_box}).

\box_if_horizontal_p:N \box_if_horizontal_p:c \box_if_horizontal:NTF \box_if_horizontal:TF

Tests if \texttt{box} is a horizontal box.

\box_if_vertical_p:N \box_if_vertical_p:c \box_if_vertical:NTF \box_if_vertical:TF

Tests if \texttt{box} is a vertical box.

5 The last box inserted

\box_set_to_last:N \box_set_to_last:c \box_gset_to_last:N \box_gset_to_last:c

Sets the \texttt{box} equal to the last item (box) added to the current partial list, removing the item from the list at the same time. When applied to the main vertical list, the \texttt{box} is always void as it is not possible to recover the last added item.

6 Constant boxes

\c_empty_box

Updated: 2012-11-04

This is a permanently empty box, which is neither set as horizontal nor vertical.

\TeXhackers note: At the \TeX level this is a void box.
7 Scratch boxes

Scratch boxes for local assignment. These are never used by the kernel code, and so are safe for use with any \texttt{L\TeX3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\begin{itemize}
\item \texttt{l_tmpa_box}
\item \texttt{l_tmpb_box}
\end{itemize}

Scratch boxes for global assignment. These are never used by the kernel code, and so are safe for use with any \texttt{L\TeX3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\begin{itemize}
\item \texttt{g_tmpa_box}
\item \texttt{g_tmpb_box}
\end{itemize}

8 Viewing box contents

\begin{itemize}
\item \texttt{\box_show:N \langle box \rangle}
\item \texttt{\box_show:c \langle box \rangle}
\end{itemize}

\texttt{\box_show:N} \langle box \rangle

Shows full details of the content of the \langle box \rangle in the terminal.

\begin{itemize}
\item \texttt{\box_show:Nnn \langle box \rangle \{\langle intexpr_1 \rangle\} \{\langle intexpr_2 \rangle\}}
\end{itemize}

\texttt{\box_show:Nnn \langle box \rangle \{\langle intexpr_1 \rangle\} \{\langle intexpr_2 \rangle\}}

Display the contents of \langle box \rangle in the terminal, showing the first \langle intexpr_1 \rangle items of the box, and descending into \langle intexpr_2 \rangle group levels.

\begin{itemize}
\item \texttt{\box_log:N \langle box \rangle}
\item \texttt{\box_log:c \langle box \rangle}
\end{itemize}

\texttt{\box_log:N} \langle box \rangle

 Writes full details of the content of the \langle box \rangle to the log.

\begin{itemize}
\item \texttt{\box_log:Nnn \langle box \rangle \{\langle intexpr_1 \rangle\} \{\langle intexpr_2 \rangle\}}
\end{itemize}

\texttt{\box_log:Nnn \langle box \rangle \{\langle intexpr_1 \rangle\} \{\langle intexpr_2 \rangle\}}

 Writes the contents of \langle box \rangle to the log, showing the first \langle intexpr_1 \rangle items of the box, and descending into \langle intexpr_2 \rangle group levels.

9 Boxes and color

All \texttt{L\TeX3} boxes are “color safe”: a color set inside the box stops applying after the end of the box has occurred.

10 Horizontal mode boxes

\begin{itemize}
\item \texttt{\hbox:n \{\langle contents \rangle\}}
\end{itemize}

\texttt{\hbox:n \{\langle contents \rangle\}}

Typesets the \langle contents \rangle into a horizontal box of natural width and then includes this box in the current list for typesetting.
\hbox_to_wd:nn \{\dimexpr\} \{\contents\}
Typesets the \{\contents\} into a horizontal box of width \{\dimexpr\} and then includes this box in the current list for typesetting.

\hbox_to_zero:n \{\contents\}
Typesets the \{\contents\} into a horizontal box of zero width and then includes this box in the current list for typesetting.

\hbox_set:Nn \hbox_set:cn \hbox_gset:Nn \hbox_gset:cn
\hbox_set_to_wd:Nnn \{\box\} \{\dimexpr\} \{\contents\}
Typesets the \{\contents\} to the width given by the \{\dimexpr\} and then stores the result inside the \{\box\}. In contrast to \hbox_set_to_wd:Nnn this function does not absorb the argument when finding the \{\content\}, and so can be used in circumstances where the \{\content\} may not be a simple argument.

\hbox_overlap_right:n \{\contents\}
Typesets the \{\contents\} into a horizontal box of zero width such that material protrudes to the right of the insertion point.

\hbox_overlap_left:n \{\contents\}
Typesets the \{\contents\} into a horizontal box of zero width such that material protrudes to the left of the insertion point.

\hbox_set:Nw \hbox_set:cw \hbox_set_end:
\hbox_set:Nw \hbox_set:cn \hbox_set_end:
\hbox_gset:Nw \hbox_gset:cn \hbox_gset_end:
\hbox_set_to_wd:Nnw \{\box\} \{\dimexpr\} \{\contents\} \hbox_set_end:
Typesets the \{\contents\} to the width given by the \{\dimexpr\} and then stores the result inside the \{\box\}. In contrast to \hbox_set_to_wd:Nnn this function does not absorb the argument when finding the \{\content\}, and so can be used in circumstances where the \{\content\} may not be a simple argument.

\hbox_unpack:N \hbox_unpack:c
Unpacks the content of the horizontal \{\box\}, retaining any stretching or shrinking applied when the \{\box\} was set.

\TeXhackers note: This is the \TeX primitive \unhcopy.
11 Vertical mode boxes

Vertical boxes inherit their baseline from their contents. The standard case is that the baseline of the box is at the same position as that of the last item added to the box. This means that the box has no depth unless the last item added to it had depth. As a result most vertical boxes have a large height value and small or zero depth. The exception are _top boxes, where the reference point is that of the first item added. These tend to have a large depth and small height, although the latter is typically non-zero.

\vbox:n {\langle contents\rangle}
Typesets the \langle contents\rangle into a vertical box of natural height and includes this box in the current list for typesetting.

\vbox_top:n {\langle contents\rangle}
Typesets the \langle contents\rangle into a vertical box of natural height and includes this box in the current list for typesetting. The baseline of the box is equal to that of the first item added to the box.

\vbox_to_ht:nn {\langle dimexpr\rangle} {\langle contents\rangle}
Typesets the \langle contents\rangle to the height given by the \langle dimexpr\rangle and then stores the result inside the \langle box\rangle.

\vbox_to_zero:n {\langle contents\rangle}
Typesets the \langle contents\rangle into a vertical box of zero height and then includes this box in the current list for typesetting.

\vbox_set:Nn \langle box\rangle {\langle contents\rangle}
Typesets the \langle contents\rangle at natural height and then stores the result inside the \langle box\rangle.

\vbox_set_top:Nn \langle box\rangle {\langle contents\rangle}
Typesets the \langle contents\rangle at natural height and then stores the result inside the \langle box\rangle. The baseline of the box is equal to that of the first item added to the box.

\vbox_set_to_ht:Nnn \langle box\rangle {\langle dimexpr\rangle} {\langle contents\rangle}
Typesets the \langle contents\rangle to the height given by the \langle dimexpr\rangle and then stores the result inside the \langle box\rangle.
\vbox_set:Nw \vbox_set:cw \vbox_set_end:
Typesets the \langle contents \rangle at natural height and then stores the result inside the \langle box \rangle. In contrast to \vbox_set:Nn this function does not absorb the argument when finding the \langle content \rangle, and so can be used in circumstances where the \langle content \rangle may not be a simple argument.

\vbox_set_to_ht:Nnw \vbox_set_to_ht:cnw \vbox_gset_to_ht:Nnw \vbox_gset_to_ht:cnw
Updated: 2017-04-05

\vbox_set_to_ht:Nw \langle box \rangle \langle\dimexpr\rangle \langle contents \rangle \vbox_set_end:
Typesets the \langle contents \rangle to the height given by the \langle dimexpr \rangle and then stores the result inside the \langle box \rangle. In contrast to \vbox_set_to_ht:Nnn this function does not absorb the argument when finding the \langle content \rangle, and so can be used in circumstances where the \langle content \rangle may not be a simple argument.

\vbox_set_split_to_ht:NNn \vbox_set_split_to_ht:(cNn|Ncn|ccn) \vbox_gset_split_to_ht:NNn \vbox_gset_split_to_ht:(cNn|Ncn|ccn)
Updated: 2018-12-29

\vbox_unpack:N \vbox_unpack:c
Unpacks the content of the vertical \langle box \rangle, retaining any stretching or shrinking applied when the \langle box \rangle was set.

\TeXhacks{This is the \TeX{} primitive \unvcopy.}

12 Using boxes efficiently

The functions above for using box contents work in exactly the same way as for any other expl3 variable. However, for efficiency reasons, it is also useful to have functions which drop box contents on use. When a box is dropped, the box becomes empty at the group level where the box was originally set rather than necessarily at the current group level. For example, with

\hbox_set:Nn \l_tmpa_box { A } \group_begin:
\hbox_set:Nn \l_tmpa_box { B } \group_begin:
\box_use_drop:N \l_tmpa_box \group_end:
\box_show:N \l_tmpa_box \group_end:
\box_show:N \l_tmpa_box

241
the first use of \texttt{\textbackslash box\_show:N} will show an entirely cleared (void) box, and the second will show the letter \emph{A} in the box.

These functions should be preferred when the content of the box is no longer required after use. Note that due to the unusual scoping behaviour of \texttt{drop} functions they may be applied to both local and global boxes: the latter will naturally be set and thus cleared at a global level.

\begin{verbatim}
\texttt{\textbackslash box\_use\_drop:N \textbackslash box\_use\_drop:c}
\end{verbatim}

Inserts the current content of the \langle box \rangle onto the current list for typesetting then drops the box content. An error is raised if the variable does not exist or if it is invalid. This function may be applied to local or global boxes.

\textbf{\LaTeXhacks note:} This is the \texttt{\textbackslash box} primitive.

\begin{verbatim}
\texttt{\textbackslash box\_set\_eq\_drop:NN \textbackslash box\_set\_eq\_drop:(cN|Nc|cc)}
\end{verbatim}

\texttt{\textbackslash box\_set\_eq\_drop:NN \texttt{(box1)} \texttt{(box2)}}

Sets the content of \texttt{(box1)} equal to that of \texttt{(box2)}, then drops \texttt{(box2)}.

\textbf{\LaTeXhacks note:} New: 2019-01-17

\begin{verbatim}
\texttt{\textbackslash box\_gset\_eq\_drop:NN \textbackslash box\_gset\_eq\_drop:(cN|Nc|cc)}
\end{verbatim}

\texttt{\textbackslash box\_gset\_eq\_drop:NN \texttt{(box1)} \texttt{(box2)}}

Sets the content of \texttt{(box1)} globally equal to that of \texttt{(box2)}, then drops \texttt{(box2)}.

\textbf{\LaTeXhacks note:} New: 2019-01-17

\begin{verbatim}
\texttt{\textbackslash hbox\_unpack\_drop:N \textbackslash hbox\_unpack\_drop:c}
\end{verbatim}

\texttt{\textbackslash hbox\_unpack\_drop:N \texttt{(box)}}

Unpacks the content of the horizontal \langle box \rangle, retaining any stretching or shrinking applied when the \langle box \rangle was set. The original \langle box \rangle is then dropped.

\textbf{\LaTeXhacks note:} This is the \LaTeX{} primitive \texttt{\textbackslash unhbox}.

\begin{verbatim}
\texttt{\textbackslash vbox\_unpack\_drop:N \textbackslash vbox\_unpack\_drop:c}
\end{verbatim}

\texttt{\textbackslash vbox\_unpack\_drop:N \texttt{(box)}}

Unpacks the content of the vertical \langle box \rangle, retaining any stretching or shrinking applied when the \langle box \rangle was set. The original \langle box \rangle is then dropped.

\textbf{\LaTeXhacks note:} This is the \LaTeX{} primitive \texttt{\textbackslash unvbox}.

\section{Affine transformations}

Affine transformations are changes which (informally) preserve straight lines. Simple translations are affine transformations, but are better handled in \LaTeX{} by doing the translation first, then inserting an unmodified box. On the other hand, rotation and resizing of boxed material can best be handled by modifying boxes. These transformations are described here.
Resizes the \langle box \rangle to fit within the given \{x-size\} (horizontally) and \{y-size\} (vertically); both of the sizes are dimension expressions. The \{y-size\} is the height only: it does not include any depth. The updated \langle box \rangle is an hbox, irrespective of the nature of the \langle box \rangle before the resizing is applied. The final size of the \langle box \rangle is the smaller of \{x-size\} and \{y-size\}, i.e. the result fits within the dimensions specified. Negative sizes cause the material in the \langle box \rangle to be reversed in direction, but the reference point of the \langle box \rangle is unchanged. Thus a negative \{y-size\} results in the \langle box \rangle having a depth dependent on the height of the original and \textit{vice versa}.

Resizes the \langle box \rangle to fit within the given \{x-size\} (horizontally) and \{y-size\} (vertically); both of the sizes are dimension expressions. The \{y-size\} is the total vertical size (height plus depth). The updated \langle box \rangle is an hbox, irrespective of the nature of the \langle box \rangle before the resizing is applied. The final size of the \langle box \rangle is the smaller of \{x-size\} and \{y-size\}, i.e. the result fits within the dimensions specified. Negative sizes cause the material in the \langle box \rangle to be reversed in direction, but the reference point of the \langle box \rangle is unchanged. Thus a negative \{y-size\} results in the \langle box \rangle having a depth dependent on the height of the original and \textit{vice versa}.

Resizes the \langle box \rangle to \{y-size\} (vertically), scaling the horizontal size by the same amount; \{y-size\} is a dimension expression. The \{y-size\} is the height only: it does not include any depth. The updated \langle box \rangle is an hbox, irrespective of the nature of the \langle box \rangle before the resizing is applied. A negative \{y-size\} causes the material in the \langle box \rangle to be reversed in direction, but the reference point of the \langle box \rangle is unchanged. Thus a negative \{y-size\} results in the \langle box \rangle having a depth dependent on the height of the original and \textit{vice versa}.
Resizes the \( \langle \text{box} \rangle \) to \( \langle y\text{-size} \rangle \) (vertically), scaling the horizontal size by the same amount: \( \langle y\text{-size} \rangle \) is a dimension expression. The \( \langle y\text{-size} \rangle \) is the total vertical size (height plus depth). The updated \( \langle \text{box} \rangle \) is an \texttt{hbox}, irrespective of the nature of the \( \langle \text{box} \rangle \) before the resizing is applied. A negative \( \langle y\text{-size} \rangle \) causes the material in the \( \langle \text{box} \rangle \) to be reversed in direction, but the reference point of the \( \langle \text{box} \rangle \) is unchanged. Thus a negative \( \langle y\text{-size} \rangle \) results in the \( \langle \text{box} \rangle \) having a depth dependent on the height of the original and \textit{vice versa}.

Resizes the \( \langle \text{box} \rangle \) to \( \langle x\text{-size} \rangle \) (horizontally), scaling the vertical size by the same amount: \( \langle x\text{-size} \rangle \) is a dimension expression. The updated \( \langle \text{box} \rangle \) is an \texttt{hbox}, irrespective of the nature of the \( \langle \text{box} \rangle \) before the resizing is applied. A negative \( \langle x\text{-size} \rangle \) causes the material in the \( \langle \text{box} \rangle \) to be reversed in direction, but the reference point of the \( \langle \text{box} \rangle \) is unchanged. Thus a negative \( \langle x\text{-size} \rangle \) results in the \( \langle \text{box} \rangle \) having a depth dependent on the height of the original and \textit{vice versa}.

Resizes the \( \langle \text{box} \rangle \) to \( \langle x\text{-size} \rangle \) (horizontally) and \( \langle y\text{-size} \rangle \) (vertically): both of the sizes are dimension expressions. The \( \langle y\text{-size} \rangle \) is the height only and does not include any depth. The updated \( \langle \text{box} \rangle \) is an \texttt{hbox}, irrespective of the nature of the \( \langle \text{box} \rangle \) before the resizing is applied. Negative sizes cause the material in the \( \langle \text{box} \rangle \) to be reversed in direction, but the reference point of the \( \langle \text{box} \rangle \) is unchanged. Thus a negative \( \langle y\text{-size} \rangle \) results in the \( \langle \text{box} \rangle \) having a depth dependent on the height of the original and \textit{vice versa}.

Resizes the \( \langle \text{box} \rangle \) to \( \langle x\text{-size} \rangle \) (horizontally) and \( \langle y\text{-size} \rangle \) (vertically): both of the sizes are dimension expressions. The \( \langle y\text{-size} \rangle \) is the total vertical size (height plus depth). The updated \( \langle \text{box} \rangle \) is an \texttt{hbox}, irrespective of the nature of the \( \langle \text{box} \rangle \) before the resizing is applied. Negative sizes cause the material in the \( \langle \text{box} \rangle \) to be reversed in direction, but the reference point of the \( \langle \text{box} \rangle \) is unchanged. Thus a negative \( \langle y\text{-size} \rangle \) results in the \( \langle \text{box} \rangle \) having a depth dependent on the height of the original and \textit{vice versa}. 
\box_rotate:Nn \box_rotate:cn
\box_grotate:Nn \box_grotate:cn
Updated: 2019-01-22
\box_scale:Nnn \box_scale:cn
\box_gscale:Nnn \box_gscale:cn
Updated: 2019-01-22

\box_rotate:Nn (box) \{angle\}
Rotates the ⟨box⟩ by ⟨angle⟩ (in degrees) anti-clockwise about its reference point. The reference point of the updated box is moved horizontally such that it is at the left side of the smallest rectangle enclosing the rotated material. The updated ⟨box⟩ is an \hbox, irrespective of the nature of the ⟨box⟩ before the rotation is applied.

\box_scale:Nnn \box_scale:cn
\box_gscale:Nnn \box_gscale:cn
Updated: 2019-01-22

14 Primitive box conditionals

\if_hbox:N * \if_hbox:N (box)
  ⟨true code⟩
\else:
  ⟨false code⟩
\fi:
Tests is ⟨box⟩ is a horizontal box.

\textbf{\TeXhackers note}: This is the \TeX primitive \ifhbox.

\if_vbox:N * \if_vbox:N (box)
  ⟨true code⟩
\else:
  ⟨false code⟩
\fi:
Tests is ⟨box⟩ is a vertical box.

\textbf{\TeXhackers note}: This is the \TeX primitive \ifvbox.

\if_box_empty:N * \if_box_empty:N (box)
  ⟨true code⟩
\else:
  ⟨false code⟩
\fi:
Tests is ⟨box⟩ is an empty (void) box.

\textbf{\TeXhackers note}: This is the \TeX primitive \ifvoid.
Part XXIX
The l3coffins package
Coffin code layer

The material in this module provides the low-level support system for coffins. For details about the design concept of a coffin, see the xcoffins module (in the l3experimental bundle).

1 Creating and initialising coffins

\coffin_new:N \coffin_new:c
Nov: 2011-08-17
Creates a new \langle coffin \rangle or raises an error if the name is already taken. The declaration is global. The \langle coffin \rangle is initially empty.

\coffin_clear:N \coffin_clear:c \coffin_gclear:N \coffin_gclear:c
Nov: 2011-08-17 Updated: 2019-01-21
Clears the content of the \langle coffin \rangle.

\coffin_set_eq:NN \coffin_set_eq:NN \coffin_gset_eq:NN \coffin_gset_eq:NN
New: 2011-08-17 Updated: 2019-01-21
Sets both the content and poles of \langle coffin_1 \rangle equal to those of \langle coffin_2 \rangle.

\coffin_if_exist_p:N \coffin_if_exist:NTF \coffin_if_exist:N \coffin_if_exist:c
Nov: 2012-06-20
Tests whether the \langle coffin \rangle is currently defined.

2 Setting coffin content and poles

\hcoffin_set:Nn \hcoffin_set:cn \hcoffin_gset:Nn \hcoffin_gset:cn
Nov: 2011-08-17 Updated: 2019-01-21
Typesets the \langle material \rangle in horizontal mode, storing the result in the \langle coffin \rangle. The standard poles for the \langle coffin \rangle are then set up based on the size of the typeset material.
\coffin_set:Nw \coffin_set:cw \coffin_set_end:
Typesets the \langle material \rangle in horizontal mode, storing the result in the \langle coffin \rangle. The standard poles for the \langle coffin \rangle are then set up based on the size of the typeset material. These functions are useful for setting the entire contents of an environment in a coffin.

\vcoffin_set:Nnn \vcoffin_set:cnn \vcoffin_gset:Nnn \vcoffin_gset:cnn
Typesets the \langle material \rangle in vertical mode constrained to the given \langle width \rangle and stores the result in the \langle coffin \rangle. The standard poles for the \langle coffin \rangle are then set up based on the size of the typeset material.

\coffin_set_horizontal_pole:Nnn \coffin_set_horizontal_pole:cnn \coffin_gset_horizontal_pole:Nnn \coffin_gset_horizontal_pole:cnn
Sets the \langle pole \rangle to run horizontally through the \langle coffin \rangle. The \langle pole \rangle is placed at the \langle offset \rangle from the bottom edge of the bounding box of the \langle coffin \rangle. The \langle offset \rangle should be given as a dimension expression.

\coffin_set_vertical_pole:Nnn \coffin_set_vertical_pole:cnn \coffin_gset_vertical_pole:Nnn \coffin_gset_vertical_pole:cnn
Sets the \langle pole \rangle to run vertically through the \langle coffin \rangle. The \langle pole \rangle is placed at the \langle offset \rangle from the left-hand edge of the bounding box of the \langle coffin \rangle. The \langle offset \rangle should be given as a dimension expression.
3 Coffin affine transformations

\coffin_resize:Nnn 
\coffin_resize:Nnn \coffin \{\textwidth\} \{\textheight\}
Resized the \coffin to \textwidth and \textheight, both of which should be given as dimension expressions.

\coffin_rotate:Nn
\coffin_rotate:cn
\coffin_grotate:Nn
\coffin_grotate:cn
Rotates the \coffin by the given \texttt{angle} (given in degrees counter-clockwise). This process rotates both the coffin content and poles. Multiple rotations do not result in the bounding box of the coffin growing unnecessarily.

\coffin_scale:Nnn
\coffin_scale:cn
\coffin_gscale:Nnn
\coffin_gscale:cn
Scales the \coffin by factors \texttt{x-scale} and \texttt{y-scale} in the horizontal and vertical directions, respectively. The two scale factors should be given as real numbers.

4 Joining and using coffins

\coffin_attach:NnnNnnnn
\coffin_attach:cn\{\texttt{coffin1}\} \{\texttt{coffin1-pole1}\} \{\texttt{coffin1-pole2}\}
\coffin_attach:cn\{\texttt{coffin2}\} \{\texttt{coffin2-pole1}\} \{\texttt{coffin2-pole2}\}
\coffin_gattach:cn\{\texttt{coffin1}\} \{\texttt{coffin1-pole1}\} \{\texttt{coffin1-pole2}\}
\coffin_gattach:cn\{\texttt{coffin2}\} \{\texttt{coffin2-pole1}\} \{\texttt{coffin2-pole2}\}
This function attaches \texttt{coffin2} to \texttt{coffin1} such that the bounding box of \texttt{coffin1} is not altered, i.e. \texttt{coffin2} can protrude outside of the bounding box of the coffin. The alignment is carried out by first calculating \texttt{handle1}, the point of intersection of \texttt{coffin1-pole1} and \texttt{coffin1-pole2}, and \texttt{handle2}, the point of intersection of \texttt{coffin2-pole1} and \texttt{coffin2-pole2}. \texttt{coffin2} is then attached to \texttt{coffin1} such that the relationship between \texttt{handle1} and \texttt{handle2} is described by the \texttt{x-offset} and \texttt{y-offset}. The two offsets should be given as dimension expressions.

\coffin_gjoin:NnnNnnnn
\coffin_gjoin:cn\{\texttt{coffin1}\} \{\texttt{coffin1-pole1}\} \{\texttt{coffin1-pole2}\}
\coffin_gjoin:cn\{\texttt{coffin2}\} \{\texttt{coffin2-pole1}\} \{\texttt{coffin2-pole2}\}
\coffin_gjoin:cn\{\texttt{coffin1}\} \{\texttt{coffin1-pole1}\} \{\texttt{coffin1-pole2}\}
\coffin_gjoin:cn\{\texttt{coffin2}\} \{\texttt{coffin2-pole1}\} \{\texttt{coffin2-pole2}\}
\coffin_gjoin:cn\{\texttt{coffin1}\} \{\texttt{coffin1-pole1}\} \{\texttt{coffin1-pole2}\}
\coffin_gjoin:cn\{\texttt{coffin2}\} \{\texttt{coffin2-pole1}\} \{\texttt{coffin2-pole2}\}
This function joins \texttt{coffin2} to \texttt{coffin1} such that the bounding box of \texttt{coffin1} may expand. The new bounding box covers the area containing the bounding boxes of the two original coffins. The alignment is carried out by first calculating \texttt{handle1}, the point of intersection of \texttt{coffin1-pole1} and \texttt{coffin1-pole2}, and \texttt{handle2}, the point of intersection of \texttt{coffin2-pole1} and \texttt{coffin2-pole2}. \texttt{coffin2} is then attached to \texttt{coffin1} such that the relationship between \texttt{handle1} and \texttt{handle2} is described by the \texttt{x-offset} and \texttt{y-offset}. The two offsets should be given as dimension expressions.
Typesetting is carried out by first calculating \( \text{handle} \), the point of intersection of \( \text{pole}_1 \) and \( \text{pole}_2 \). The coffin is then typeset in horizontal mode such that the relationship between the current reference point in the document and the \( \text{handle} \) is described by the \( \text{x-offset} \) and \( \text{y-offset} \). The two offsets should be given as dimension expressions. Typesetting a coffin is therefore analogous to carrying out an alignment where the “parent” coffin is the current insertion point.

5 Measuring coffins

\[ \text{coffin}\_dp:N \{ \text{coffin} \} \]
\[ \text{coffin}\_dp:C \]
Calculates the depth (below the baseline) of the \( \text{coffin} \) in a form suitable for use in a \( \text{dimension expression} \).

\[ \text{coffin}\_ht:N \{ \text{coffin} \} \]
\[ \text{coffin}\_ht:C \]
Calculates the height (above the baseline) of the \( \text{coffin} \) in a form suitable for use in a \( \text{dimension expression} \).

\[ \text{coffin}\_wd:N \{ \text{coffin} \} \]
\[ \text{coffin}\_wd:C \]
Calculates the width of the \( \text{coffin} \) in a form suitable for use in a \( \text{dimension expression} \).

6 Coffin diagnostics

\[ \text{coffin}\_display\_handles:Nn \{ \text{coffin} \} \{\{\text{color}\}\} \]
This function first calculates the intersections between all of the \( \text{poles} \) of the \( \text{coffin} \) to give a set of \( \text{handles} \). It then prints the \( \text{coffin} \) at the current location in the source, with the position of the \( \text{handles} \) marked on the coffin. The \( \text{handles} \) are labelled as part of this process: the locations of the \( \text{handles} \) and the labels are both printed in the \( \text{color} \) specified.

\[ \text{coffin}\_mark\_handle:Nnnn \{ \text{coffin} \} \{\{\text{pole}_1\}\} \{\{\text{pole}_2\}\} \{\{\text{color}\}\} \]
This function first calculates the \( \text{handle} \) for the \( \text{coffin} \) as defined by the intersection of \( \text{pole}_1 \) and \( \text{pole}_2 \). It then marks the position of the \( \text{handle} \) on the \( \text{coffin} \). The \( \text{handle} \) are labelled as part of this process: the location of the \( \text{handle} \) and the label are both printed in the \( \text{color} \) specified.

\[ \text{coffin}\_show\_structure:N \{ \text{coffin} \} \]
This function shows the structural information about the \( \text{coffin} \) in the terminal. The width, height and depth of the typeset material are given, along with the location of all of the poles of the coffin.

Notice that the poles of a coffin are defined by four values: the \( x \) and \( y \) co-ordinates of a point that the pole passes through and the \( x \)- and \( y \)-components of a vector denoting the direction of the pole. It is the ratio between the later, rather than the absolute values, which determines the direction of the pole.
This function writes the structural information about the \textit{coffin} in the log file. See also \texttt{coffin\_show\_structure} which displays the result in the terminal.

## 7 Constants and variables

\texttt{c\_empty\_coffin}

A permanently empty coffin.

\texttt{l\_tmpa\_coffin} \texttt{l\_tmpb\_coffin}

Scratch coffins for local assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX\-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

\texttt{g\_tmpa\_coffin} \texttt{g\_tmpb\_coffin}

Scratch coffins for global assignment. These are never used by the kernel code, and so are safe for use with any \LaTeX\-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
Part XXX

The \texttt{l3color-base} package

Color support

This module provides support for color in \TeX3. At present, the material here is mainly intended to support a small number of low-level requirements in other \texttt{l3kernel} modules.

1 Color in boxes

Controlling the color of text in boxes requires a small number of control functions, so that the boxed material uses the color at the point where it is set, rather than where it is used.

\begin{verbatim}
\color_group_begin:
\color_group_end:
\end{verbatim}

New: 2011-09-03

\begin{verbatim}
\color_ensure_current:
\end{verbatim}

New: 2011-09-03

\begin{verbatim}
\color_group_begin:
...\color_group_end:
\end{verbatim}

Creates a color group: one used to “trap” color settings.

\begin{verbatim}
\color_ensure_current:
\end{verbatim}

Ensures that material inside a box uses the foreground color at the point where the box is set, rather than that in force when the box is used. This function should usually be used within a \texttt{\color_group_begin}: \ldots\texttt{\color_group_end}: group.
Part XXXI

The \texttt{l3luatex} package:
Lua\TeX{}-specific functions

The Lua\TeX{} engine provides access to the Lua programming language, and with it access
to the “internals” of \TeX{}. In order to use this within the framework provided here, a
family of functions is available. When used with pdf\TeX, p\TeX{}, up\TeX{} or X\TeX{} these
raise an error: use \texttt{\sys_if_engine_luatex:T} to avoid this. Details on using Lua with
the Lua\TeX{} engine are given in the Lua\TeX{} manual.

1 Breaking out to Lua

\texttt{\lua_now:n \{token list\}}

The \langle token list \rangle is first tokenized by \TeX{}, which includes converting line ends to spaces in
the usual \TeX{} manner and which respects currently-applicable \TeX{} category codes. The
resulting \langle Lua input \rangle is passed to the Lua interpreter for processing. Each \texttt{\lua_now:n}
block is treated by Lua as a separate chunk. The Lua interpreter executes the \langle Lua input \rangle
immediately, and in an expandable manner.

\texttt{\lua_now:e} is a macro wrapper around \texttt{\directlua}: when Lua\TeX{} is
in use two expansions are required to yield the result of the Lua code.

\texttt{\lua_shipout:n \{token list\}}

The \langle token list \rangle is first tokenized by \TeX{}, which includes converting line ends to spaces in
the usual \TeX{} manner and which respects currently-applicable \TeX{} category codes. The
resulting \langle Lua input \rangle is passed to the Lua interpreter when the current page is finalised
(i.e. at shipout). Each \texttt{\lua_shipout:n} block is treated by Lua as a separate chunk.
The Lua interpreter will execute the \langle Lua input \rangle during the page-building routine: no
\TeX{} expansion of the \langle Lua input \rangle will occur at this stage.

In the case of the \texttt{\lua_shipout_e:n} version the input is fully expanded by \TeX{} in
an e-type manner during the shipout operation.

\texttt{\lua_shipout_e:n} is a macro wrapper around \texttt{\luaescapestring}: when Lua\TeX{} is
in use two expansions are required to yield the result of the Lua code.

\texttt{\lua_escape:n \{token list\}}

Converts the \langle token list \rangle such that it can safely be passed to Lua: embedded backslashes,
double and single quotes, and newlines and carriage returns are escaped. This is done by
prepending an extra token consisting of a backslash with category code 12, and for the
line endings, converting them to \texttt{\n} and \texttt{\r}, respectively.

\texttt{\lua_escape:e} is a macro wrapper around \texttt{\luaescapestring}: when Lua\TeX{} is
in use two expansions are required to yield the result of the Lua code.
2 Lua interfaces

As well as interfaces for \TeX, there are a small number of Lua functions provided here.

\begin{verbatim}
\texttt{\textbackslash l3kernel.charcat(⟨charcode⟩, ⟨catcode⟩)}
\end{verbatim}

Constructs a character of ⟨charcode⟩ and ⟨catcode⟩ and returns the result to \TeX.

\begin{verbatim}
\texttt{\textbackslash l3kernel.elapsedtime()}\texttt{\textbackslash \textbackslash l3kernel.filedump(⟨file⟩,⟨offset⟩,⟨length⟩)}\texttt{\textbackslash \textbackslash l3kernel.filemdfivesum(⟨file⟩)}\texttt{\textbackslash \textbackslash l3kernel.filemoddate(⟨file⟩)}\texttt{\textbackslash \textbackslash l3kernel.filesize(⟨file⟩)}\texttt{\textbackslash \textbackslash l3kernel.resettimer()}\texttt{\textbackslash \textbackslash l3kernel.shellescape(⟨cmd⟩)}\texttt{\textbackslash \textbackslash l3kernel.strcmp(⟨str one⟩, ⟨str two⟩)}
\end{verbatim}

<table>
<thead>
<tr>
<th>\texttt{l3kernel}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{l3kernel.charcat(⟨charcode⟩, ⟨catcode⟩)}</td>
</tr>
<tr>
<td>Constructs a character of ⟨charcode⟩ and ⟨catcode⟩ and returns the result to \TeX.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.elapsedtime()}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns the CPU time in \textit{scaled seconds} since the start of the \TeX run or since \texttt{l3kernel.resettimer} was issued. This only measures the time used by the CPU, not the real time, e.g., waiting for user input.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.filedump(⟨file⟩,⟨offset⟩,⟨length⟩)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns the uppercase hexadecimal representation of the content of the ⟨file⟩ read as bytes. If the ⟨length⟩ is given, only this part of the file is returned; similarly, one may specify the ⟨offset⟩ from the start of the file. If the ⟨length⟩ is not given, the entire file is read starting at the ⟨offset⟩.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.filemdfivesum(⟨file⟩)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns the \textit{of the MD5 sum of the file contents read as bytes; note that the result will depend on the nature of the line endings used in the file, in contrast to normal \TeX behaviour. If the ⟨file⟩ is not found, nothing is returned with \textit{no error raised}.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.filemoddate(⟨file⟩)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns the \textit{of the date/time of last modification of the ⟨file⟩ in the format}</td>
</tr>
<tr>
<td>\texttt{D:⟨year⟩⟨month⟩⟨day⟩⟨hour⟩⟨minute⟩⟨second⟩⟨offset⟩}</td>
</tr>
<tr>
<td>where the latter may be \texttt{Z} (UTC) or ⟨plus-minus⟩\texttt{⟨hours⟩‘⟨minutes⟩‘}. If the ⟨file⟩ is not found, nothing is returned with \textit{no error raised}.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.filesize(⟨file⟩)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns the size of the ⟨file⟩ in bytes. If the ⟨file⟩ is not found, nothing is returned with \textit{no error raised}.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.resettimer()}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resets the timer used by \texttt{l3kernel.elapsedtime}.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.shellescape(⟨cmd⟩)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executes the ⟨cmd⟩ and prints to the log as for pdf\TeX.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\texttt{l3kernel.strcmp(⟨str one⟩, ⟨str two⟩)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compares the two strings and returns 0 to \TeX if the two are identical.</td>
</tr>
</tbody>
</table>
Part XXXII

The l3unicode package: Unicode support functions

This module provides Unicode-specific functions along with loading data from a range of Unicode Consortium files. At present, it provides no public functions.
Part XXXIII

The l3text package: text processing

1 l3text documentation

This module deals with manipulation of (formatted) text; such material is comprised of a restricted set of token list content. The functions provided here concern conversion of textual content for example in case changing, generation of bookmarks and extraction to tags. All of the major functions operate by expansion. Begin-group and end-group tokens in the ⟨text⟩ are normalized and become { and }, respectively.

1.1 Expanding text

\text_expand:n \{⟨text⟩\}

Takes user input ⟨text⟩ and expands the content. Protected commands (typically formatting) are left in place, and no processing takes place of math mode material (as delimited by pairs given in \l_text_math_delims_tl or as the argument to commands listed in \l_text_math_arg_tl). Commands which are neither engine- nor \LaTeX{} protected are expanded exhaustively. Any commands listed in \l_text_expand_exclude_tl, \l_text_accent_tl and \l_text_letterlike_tl are excluded from expansion.

\text_declare_expand_equivalent:Nn \{⟨cmd⟩\} ⟨replacement⟩

\text_declare_expand_equivalent:cn

Declares that the ⟨replacement⟩ tokens should be used whenever the ⟨cmd⟩ (a single token) is encountered. The ⟨replacement⟩ tokens should be expandable.
1.2 Case changing

\text_lowercase:n \{\text{tokens}\}
\text_uppercase:n \{\text{tokens}\}
\text_titlecase:n \{\text{language}\} \{\text{tokens}\}
\text_titlecase_first:n
\text_lowercase:nn
\text_uppercase:nn
\text_titlecase:nn
\text_titlecase_first:nn

Takes user input (\text{text}) first applies \text{\textunderscore expand}, then transforms the case of character tokens as specified by the function name. The category code of letters are not changed by this process (at least where they can be represented by the engine as a single token: 8-bit engines may require active characters).

Upper- and lowercase have the obvious meanings. Titlecasing may be regarded informally as converting the first character of the (\text{tokens}) to uppercase and the rest to lowercase. However, the process is more complex than this as there are some situations where a single lowercase character maps to a special form, for example \textit{i}j in Dutch which becomes \textit{IJ}. The \text{\textunderscore titlecase\_first} variant does not attempt any case changing at all after the first letter has been processed.

Importantly, notice that these functions are intended for working with user text for typesetting. For case changing programmatic data see the \l3str module and discussion there of \textit{\textunderscore str\_lowercase:n}, \textit{\textunderscore str\_uppercase:n} and \textit{\textunderscore str\_foldcase:n}.

Case changing does not take place within math mode material so for example

\textit{\textunderscore uppercase:n} \{ \textit{Some-text}=$x = mx + c$ with {Braces} \}

becomes

\textit{SOME TEXT} $y = mx + c$ WITH \{BRACES\}

The arguments of commands listed in \texttt{\textunderscore l\_text\_case\_exclude\_arg\_tl} are excluded from case changing; the latter are entirely non-textual content (such as labels).

As is generally true for \texttt{expl3}, these functions are designed to work with Unicode input only. As such, UTF-8 input is assumed for all engines. When used with Xe\TeX{} or Lua\TeX{} a full range of Unicode transformations are enabled. Specifically, the standard mappings here follow those defined by the Unicode Consortium in UnicodeData.txt and SpecialCasing.txt. In the case of 8-bit engines, mappings are provided for characters which can be represented in output typeset using the T1, T2 and LGR font encodings. Thus for example \textit{ä} can be case-changed using pdf\TeX. For \texttt{pdflatex} only the ASCII range is covered as the engine treats input outside of this range as east Asian.

Language-sensitive conversions are enabled using the (\text{language}) argument, and follow Unicode Consortium guidelines. Currently, the languages recognised for special handling are as follows.

- **Azeri and Turkish** (az and tr). The case pairs I/i-dotless and I-dot/i are activated for these languages. The combining dot mark is removed when lowercasing I-dot and introduced when upper casing i-dotless.

- **German** (de-alt). An alternative mapping for German in which the lowercase \textit{Eszet} maps to a \textit{großes Eszet}. Since there is a T1 slot for the \textit{großes Eszet} in T1, this tailoring is available with pdf\TeX{} as well as in the Unicode \TeX{} engines.

- **Greek** (el). Removes accents from Greek letters when uppcasing; titlecasing leaves accents in place.

- **Lithuanian** (lt). The lowercase letters i and j should retain a dot above when the accents grave, acute or tilde are present. This is implemented for lowercasing of the relevant uppercase letters both when input as single Unicode codepoints and when using combining accents. The combining dot is removed when uppcasing in these cases. Note that only the accents used in Lithuanian are covered: the behaviour of other accents are not modified.

- **Dutch** (nl). Capitalisation of \textit{i}j at the beginning of titlecased input produces \textit{IJ} rather than \textit{I}j. The output retains two separate letters, thus this transformation is available using pdf\TeX{}}.
1.3 Removing formatting from text

\text_purify:n \{⟨text⟩\}

Takes user input ⟨text⟩ and expands as described for \text_expand:n, then removes all functions from the resulting text. Math mode material (as delimited by pairs given in \l_text_math_delims_tl or as the argument to commands listed in \l_text_math_arg_tl) is left contained in a pair of $ delimiters. Non-expandable functions present in the ⟨text⟩ must either have a defined equivalent (see \text_declare_purify_equivalent:Nn) or will be removed from the result. Implicit tokens are converted to their explicit equivalent.

\text_declare_purify_equivalent:Nn \text_declare_purify_equivalent:Nn (cmd) {⟨replacement⟩}
\text_declare_purify_equivalent:Nx

Declares that the ⟨replacement⟩ tokens should be used whenever the ⟨cmd⟩ (a single token) is encountered. The ⟨replacement⟩ tokens should be expandable.

1.4 Control variables

\l_text_accents_tl Lists commands which represent accents, and which are left unchanged by expansion. (Defined only for the \LaTeX2ε package.)

\l_text_letterlike_tl Lists commands which represent letters; these are left unchanged by expansion. (Defined only for the \LaTeX2ε package.)

\l_text_math_arg_tl Lists commands present in the ⟨text⟩ where the argument of the command should be treated as math mode material. The treatment here is similar to \l_text_math_delims_tl but for a command rather than paired delimiters.

\l_text_math_delims_tl Lists pairs of tokens which delimit (in-line) math mode content; such content may be excluded from processing.

\l_text_expand_exclude_tl Lists commands which are excluded from expansion.

New: 2020-03-05

\text_declare_purify_equivalent:Nn
\text_declare_purify_equivalent:Nx

Declares that the ⟨replacement⟩ tokens should be used whenever the ⟨cmd⟩ (a single token) is encountered. The ⟨replacement⟩ tokens should be expandable.
Part XXXIV

The \texttt{l3legacy} package

Interfaces to legacy concepts

There are a small number of \TeX{} or \LaTeX{} concepts which are not used in \texttt{expl3} code but which need to be manipulated when working as a \LaTeX{} package. To allow these to be integrated cleanly into \texttt{expl3} code, a set of legacy interfaces are provided here.

\begin{verbatim}
\legacy_if:nTF \{\langle name \rangle\} \{\langle true code \rangle\} \{\langle false code \rangle\}
\legacy_if_p:n \legacy_if:nTF
\end{verbatim}

Tests if the \LaTeX/\plain \TeX{} conditional (generated by \texttt{\newif}) if \texttt{true} or \texttt{false} and branches accordingly. The \langle name \rangle of the conditional should \textit{omit} the leading \texttt{if}. 
Part XXXV

The l3candidates package

Experimental additions to l3kernel

1 Important notice

This module provides a space in which functions can be added to l3kernel (expl3) while still being experimental.

As such, the functions here may not remain in their current form, or indeed at all, in l3kernel in the future.

In contrast to the material in l3experimental, the functions here are all small additions to the kernel. We encourage programmers to test them out and report back on the LaTeX-L mailing list.

Thus, if you intend to use any of these functions from the candidate module in a public package offered to others for productive use (e.g., being placed on CTAN) please consider the following points carefully:

• Be prepared that your public packages might require updating when such functions are being finalized.

• Consider informing us that you use a particular function in your public package, e.g., by discussing this on the LaTeX-L mailing list. This way it becomes easier to coordinate any updates necessary without issues for the users of your package.

• Discussing and understanding use cases for a particular addition or concept also helps to ensure that we provide the right interfaces in the final version so please give us feedback if you consider a certain candidate function useful (or not).

We only add functions in this space if we consider them being serious candidates for a final inclusion into the kernel. However, real use sometimes leads to better ideas, so functions from this module are not necessarily stable and we may have to adjust them!

2 Additions to l3box

2.1 Viewing part of a box

\box_clip:N \box_clip:c
\box_gclip:N \box_gclip:c

Clips the \texttt{\langle box\rangle} in the output so that only material inside the bounding box is displayed in the output. The updated \texttt{\langle box\rangle} is an hbox, irrespective of the nature of the \texttt{\langle box\rangle} before the clipping is applied.

These functions require the L\TeX\!X\!3 native drivers: they do not work with the L\TeX\!X\!2\!\epsilon\! graphics drivers!

\TeX\!X\!hackers note: Clipping is implemented by the driver, and as such the full content of the box is placed in the output file. Thus clipping does not remove any information from the raw output, and hidden material can therefore be viewed by direct examination of the file.
\box_set_trim:Nnnnn \box_set_trim:cnnnn \box_gset_trim:Nnnnn \box_gset_trim:cnnnn

Adjusts the bounding box of the \langle box \rangle \langle left \rangle from the left-hand edge of the bounding box, \langle right \rangle from the right-hand edge and so forth. All adjustments are \langle dimension expressions \rangle. Material outside of the bounding box is still displayed in the output unless \box_clip:N is subsequently applied. The updated \langle box \rangle is an hbox, irrespective of the nature of the \langle box \rangle before the trim operation is applied. The behavior of the operation where the trims requested is greater than the size of the box is undefined.

\box_set_viewport:Nnnnn \box_set_viewport:cnnnn \box_gset_viewport:Nnnnn \box_gset_viewport:cnnnn

Adjusts the bounding box of the \langle box \rangle such that it has lower-left co-ordinates ((llx), (lly)) and upper-right co-ordinates ((urx), (ury)). All four co-ordinate positions are \langle dimension expressions \rangle. Material outside of the bounding box is still displayed in the output unless \box_clip:N is subsequently applied. The updated \langle box \rangle is an hbox, irrespective of the nature of the \langle box \rangle before the viewport operation is applied.

3 Additions to l3expan

\exp_args_generate:n

\exp_args_generate:n \langle \{variant argument specifiers\} \rangle

Defines \exp_args:N\langle variant \rangle functions for each \langle variant \rangle given in the comma list \{\langle variant argument specifiers \rangle\}. Each \langle variant \rangle should consist of the letters N, c, n, V, v, o, f, e, x, p and the resulting function is protected if the letter x appears in the \langle variant \rangle. This is only useful for cases where \cs_generate_variant:Nn is not applicable.

4 Additions to l3fp

\fp_if_nan:n \fp_if_nan:n \langle \{fexpr\} \rangle

Evaluates the \(fexpr\) and tests whether the result is exactly NaN. The test returns false for any other result, even a tuple containing NaN.

5 Additions to l3file

\iow_allow_break:

\iow_allow_break:

In the first argument of \iow_wrap:nnnN (for instance in messages), inserts a break-point that allows a line break. In other words this is a zero-width breaking space.

261
\ior_get_term:nN \ior_str_get_term:nN

New: 2019-03-23

\ior_get_term:nN \ior_str_get_term:nN \langle token list variable \rangle

Function that reads one or more lines (until an equal number of left and right braces are found) from the terminal and stores the result locally in the \langle token list \rangle variable. Tokenization occurs as described for \ior_get:NN or \ior_str_get:NN, respectively. When the \langle prompt \rangle is empty, \TeX{} will wait for input without any other indication: typically the programmer will have provided a suitable text using e.g. \iorw_term:n. Where the \langle prompt \rangle is given, it will appear in the terminal followed by an =, e.g.

\begin{verbatim}
prompt=
\end{verbatim}

\ior_shell_open:nN \ior_shell_open:n \langle stream \rangle \{ \langle shell command \rangle \}

Opens the pseudo-file created by the output of the \langle shell command \rangle for reading using \langle stream \rangle as the control sequence for access. If the \langle stream \rangle was already open it is closed before the new operation begins. The \langle stream \rangle is available for access immediately and will remain allocated to \langle shell command \rangle until a \ior_close:N instruction is given or the \TeX{} run ends. If piped system calls are disabled an error is raised.

For details of handling of the \langle shell command \rangle, see \sys_get_shell:nnN(TF).

6 Additions to l3flag

\flag_raise_if_clear:n \star \flag_raise_if_clear:n \langle flag name \rangle

Ensures the \langle flag \rangle is raised by making its height at least 1, locally.

7 Additions to l3intarray

\intarray_gset_rand:Nnn \intarray_gset_rand:cn \intarray_gset_rand:Nn \intarray_gset_rand:cn

New: 2018-05-05

\intarray_gset_rand:Nnn \langle intarray \rangle \{ \langle minimum \rangle \} \{ \langle maximum \rangle \}
\intarray_gset_rand:cn \langle intarray \rangle \{ \langle maximum \rangle \}

Evaluates the integer expressions \langle minimum \rangle and \langle maximum \rangle then sets each entry (independently) of the \langle integer array variable \rangle to a pseudo-random number between the two (with bounds included). If the absolute value of either bound is bigger than \(2^{30} - 1\), an error occurs. Entries are generated in the same way as repeated calls to \int_rand:nn or \int_rand:n respectively, in particular for the second function the \langle minimum \rangle is 1. Assignments are always global. This is not available in older versions of \XeTeX{}.

7.1 Working with contents of integer arrays

\intarray_to_clist:N \star \intarray_to_clist:N \langle intarray \rangle

New: 2018-05-04

Converts the \langle intarray \rangle to integer denotations separated by commas. All tokens have category code other. If the \langle intarray \rangle has no entry the result is empty; otherwise the result has one fewer comma than the number of items.
8 Additions to l3msg

In very rare cases it may be necessary to produce errors in an expansion-only context. The functions in this section should only be used if there is no alternative approach using \msg_error:nnnnnn or other non-expandable commands from the previous section. Despite having a similar interface as non-expandable messages, expandable errors must be handled internally very differently from normal error messages, as none of the tools to print to the terminal or the log file are expandable. As a result, short-hands such as \{ or \ do not work, and messages must be very short (with default settings, they are truncated after approximately 50 characters). It is advisable to ensure that the message is understandable even when truncated, by putting the most important information up front. Another particularity of expandable messages is that they cannot be redirected or turned off by the user.

\msg_expansible_error:nnnnnn * \msg_expansible_error:nnnnn \{\module\} \{\message\} \{\arg one\} \{\arg two\} \{\arg three\} \{\arg four\}

Issues an “Undefined error” message from T\TeX\ itself using the undefined control sequence \::error then prints “! \{\module\}: ”\{\error message\}, which should be short. With default settings, anything beyond approximately 60 characters long (or bytes in some engines) is cropped. A leading space might be removed as well.

\msg_show_eval:Nn \msg_log_eval:Nn

\msg_show:nnnnnn \msg_show:nnnxxx \msg_show:nnnnn \msg_show:nnxxx \msg_show:nnnx \msg_show:nnx \msg_show:nn

Issues \{\module\} information \{\message\}, passing \{\arg one\} to \{\arg four\} to the text-creating functions. The information text is shown on the terminal and the T\TeX\ run is interrupted in a manner similar to \tl_show:n. This is used in conjunction with \msg_show_item:n and similar functions to print complex variable contents completely. If the formatted text does not contain \textgreater at the start of a line, an additional line \textgreater will be put at the end. In addition, a final period is added if not present.

New: 2017-12-04
Updated: 2019-02-28

New: 2015-08-06
Updated: 2019-02-28

263
Used in the text of messages for \texttt{msg\_show::nnxxxx} to show or log a list of items or key–value pairs. The one-argument functions are used for sequences, clist or token lists and the others for property lists. These functions turn their arguments to strings.

\section{Additions to \texttt{l3prg}}

\begin{verbatim}
\bool_set_inverse:N \bool_set_inverse:N \bool_set_inverse:c \bool_set_inverse:N \bool_set_inverse:c
\end{verbatim}

\texttt{\bool_set_inverse:N} \texttt{\bool_set_inverse:c}

Toggles the \texttt{\langle boolean\rangle} from \texttt{true} to \texttt{false} and conversely: sets it to the inverse of its current value.

\begin{verbatim}
\bool_case_true:nTF \bool_case_true:nTF \bool_case_false:nTF \bool_case_false:nTF
\end{verbatim}

\texttt{\bool_case_true:nTF}

\texttt{\bool_case_false:nTF}

Evaluates in turn each of the \texttt{\langle boolean expression cases\rangle} until the first one that evaluates to \texttt{true} or to \texttt{false}, for \texttt{\bool_case_true:n} and \texttt{\bool_case_false:n}, respectively. The \texttt{\langle code\rangle} associated to this first case is left in the input stream, followed by the \texttt{\langle true code\rangle}, and other cases are discarded. If none of the cases match then only the \texttt{\langle false code\rangle} is inserted. The functions \texttt{\bool_case_true:n} and \texttt{\bool_case_false:n}, which do nothing if there is no match, are also available. For example

\begin{verbatim}
\bool_case_true:nF
{ \dim_compare_p:n { \l__mypkg_wd_dim <= 10pt } } { Fits }
\int_compare_p:n { \l__mypkg_total_int >= 10 } } { Many }
\l__mypkg_special_bool } { Special }
\end{verbatim}

leaves “Fits” or “Many” or “Special” or “No idea!” in the input stream, in a way similar to some other language’s “if \texttt{\ldots elseif \ldots elseif \ldots else \ldots}.”
10 Additions to \texttt{l3prop}

\begin{itemize}
\item \texttt{\prop\_rand\_key\_value:N} \hspace{1em} \texttt{\prop\_rand\_key\_value:c}
\end{itemize}

Selects a pseudo-random key–value pair from the \texttt{property list} and returns \{\texttt{key}\\} and \{\texttt{value}\}. If the \texttt{property list} is empty the result is empty. This is not available in older versions of XeT\TeX.

\textbf{T\TeXhackers note:} The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp\_not:n}), which means that the \texttt{\value} does not expand further when appearing in a \texttt{x}-type argument expansion.

11 Additions to \texttt{l3seq}

\begin{itemize}
\item \texttt{\seq\_map\_thread\_function:NNN} \hspace{1em} \texttt{\seq\_map\_thread\_function:NNN (seq1) (seq2) (function)}
\item \texttt{\seq\_set\_filter:NNn} \hspace{1em} \texttt{\seq\_gset\_filter:NNn}
\item \texttt{\seq\_set\_map:NNn} \hspace{1em} \texttt{\seq\_gset\_map:NNn}
\end{itemize}

\texttt{\seq\_map\_thread\_function:NNN} \hspace{1em} \texttt{\seq\_map\_thread\_function:NNN (seq1) (seq2) (function)}

Applies \texttt{\function} to every pair of items \texttt{\seq1-item}–\texttt{\seq2-item} from the two sequences, returning items from both sequences from left to right. The \texttt{\function} receives two \texttt{n}-type arguments for each iteration. The mapping terminates when the end of either sequence is reached \textit{(i.e.} whichever sequence has fewer items determines how many iterations occur).

\texttt{\seq\_set\_filter:NNn} \hspace{1em} \texttt{\seq\_gset\_filter:NNn}

\texttt{\seq\_set\_filter:NNn (sequence1) (sequence2) \{\texttt{inline boolean}\}}

Evaluates the \texttt{\inline boolean} for every \texttt{\item} stored within the \texttt{\sequence2}. The \texttt{\inline boolean} receives the \texttt{\item} as \texttt{\#1}. The sequence of all \texttt{\items} for which the \texttt{\inline boolean} evaluated to \texttt{true} is assigned to \texttt{\sequence1}.

\textbf{T\TeXhackers note:} Contrarily to other mapping functions, \texttt{\seq\_map\_break}: cannot be used in this function, and would lead to low-level T\TeX errors.

\texttt{\seq\_set\_map:NNn} \hspace{1em} \texttt{\seq\_gset\_map:NNn}

\texttt{\seq\_set\_map:NNn (sequence1) (sequence2) \{\texttt{inline function}\}}

Applies \texttt{\inline function} to every \texttt{\item} stored within the \texttt{\sequence2}. The \texttt{\inline function} should consist of code which will receive the \texttt{\item} as \texttt{\#1}. The sequence resulting from \texttt{x}-expanding \texttt{\inline function} applied to each \texttt{\item} is assigned to \texttt{\sequence1}. As such, the code in \texttt{\inline function} should be expandable.

\textbf{T\TeXhackers note:} Contrarily to other mapping functions, \texttt{\seq\_map\_break}: cannot be used in this function, and would lead to low-level T\TeX errors.
\seq_set_from_function:Nn \seq_gset_from_function:Nn \seq_set_from_inline_x:Nnn \seq_gset_from_inline_x:Nnn
\seq_indexed_map_function:NN \seq_indexed_map_inline:Nn

Sets the \texttt{\seq} equal to a sequence whose items are obtained by \texttt{x}-expanding \texttt{\loop code} \texttt{\function}. This expansion must result in successive calls to the \texttt{\function} with no nonexpandable tokens in between. More precisely the \texttt{\function} is replaced by a wrapper function that inserts the appropriate separators between items in the sequence. The \texttt{\loop code} must be expandable; it can be for example \texttt{\tl_map_function:NN \tl var} or \texttt{\clist_map_function:nN \{clist\}} or \texttt{\int_step_function:nnnN \{initial value\} \{step\} \{final value\}}.

\seq_indexed_map_function:NN \seq_indexed_map_inline:Nn

Sets the \texttt{\seq} equal to a sequence whose items are obtained by \texttt{x}-expanding \texttt{\loop code} applied to a \texttt{\function} derived from the \texttt{\inline code}. A \texttt{\function} is defined, that takes one argument, \texttt{x}-expands the \texttt{\inline code} with that argument as \texttt{#1}, then adds appropriate separators to turn the result into an item of the sequence. The \texttt{x}-expansion of \texttt{\loop code} \texttt{\function} must result in successive calls to the \texttt{\function} with no nonexpandable tokens in between. The \texttt{\loop code} must be expandable; it can be for example \texttt{\tl_map_function:NN \tl var} or \texttt{\clist_map_function:nN \{clist\}} or \texttt{\int_step_function:nnnN \{initial value\} \{step\} \{final value\}}, but not the analogous “inline” mappings.

\seq_indexed_map_function:NN \seq_indexed_map_inline:Nn

Applies \texttt{\function} to every entry in the \texttt{\sequence variable}. The \texttt{\function} should have signature \texttt{:nn}. It receives two arguments for each iteration: the \texttt{\index} (namely 1 for the first entry, then 2 and so on) and the \texttt{\item}.

\seq_indexed_map_inline:Nn

Applies \texttt{\inline function} to every entry in the \texttt{\sequence variable}. The \texttt{\inline function} should consist of code which receives the \texttt{\index} (namely 1 for the first entry, then 2 and so on) as \texttt{#1} and the \texttt{\item} as \texttt{#2}.
12 Additions to \texttt{l3sys}

\texttt{\c_sys_engine_version_str}

The version string of the current engine, in the same form as given in the banner issued when running a job. For \texttt{pdf\TeX} and \texttt{Lua\TeX} this is of the form

\(\langle \text{major} \rangle.\langle \text{minor} \rangle.\langle \text{revision} \rangle\)

For \texttt{Xe\TeX}, the form is

\(\langle \text{major} \rangle.\langle \text{minor} \rangle\)

For \texttt{p\TeX} and \texttt{up\TeX}, only releases since \TeX\ Live 2018 make the data available, and the form is more complex, as it comprises the \texttt{p\TeX} version, the \texttt{up\TeX} version and the e-\texttt{p\TeX} version.

\(p\langle \text{major} \rangle.\langle \text{revision} \rangle-u\langle \text{major} \rangle.\langle \text{minor} \rangle-\langle \text{e\TeX} \rangle\)

where the \texttt{u} part is only present for \texttt{up\TeX}.

\texttt{\sys_if_rand_exist_p: ∗ \sys_if_rand_exist:TF ∗}

Tests if the engine has a pseudo-random number generator. Currently this is the case in \texttt{pdf\TeX}, \texttt{Lua\TeX}, \texttt{p\TeX}, \texttt{up\TeX} and recent releases of \texttt{Xe\TeX}.

\texttt{\sys_if_rand_exist_p: ∗ \sys_if_rand_exist:TF \{(true code\}) \{(false code\})}
Leaves in the input stream the items from the \langle start index \rangle to the \langle end index \rangle inclusive, using the same indexing as \texttt{l3tl\_range:nnn}. Spaces are ignored. Regardless of whether items appear with or without braces in the \langle token list \rangle, the \texttt{l3tl\_range\_braced:nnn} function wraps each item in braces, while \texttt{l3tl\_range\_unbraced:nnn} does not (overall it removes an outer set of braces). For instance,

\begin{verbatim}
\io Term:x { \tl_range_braced:nnn { abcd\{-e\}f } { 2 } { 5 } }
\io Term:x { \tl_range_braced:nnn { abcd\{-e\}f } { -4 } { -1 } }
\io Term:x { \tl_range_braced:nnn { abcd\{-e\}f } { -2 } { -1 } }
\io Term:x { \tl_range_braced:nnn { abcd\{-e\}f } { 0 } { -1 } }
\end{verbatim}

prints \texttt{\{b\}c\{d\}\{e\}}{f}, \texttt{\{c\}d\{e\}}{f}, \{f\}, and an empty line to the terminal, while

\begin{verbatim}
\io Term:x { \tl_range_unbraced:nnn { abcd\{-e\}f } { 2 } { 5 } }
\io Term:x { \tl_range_unbraced:nnn { abcd\{-e\}f } { -4 } { -1 } }
\io Term:x { \tl_range_unbraced:nnn { abcd\{-e\}f } { -2 } { -1 } }
\io Term:x { \tl_range_unbraced:nnn { abcd\{-e\}f } { 0 } { -1 } }
\end{verbatim}

prints \texttt{bcde\{}{f}, \texttt{cde\}}{f}, \texttt{e\}}{f}, and an empty line to the terminal. Because braces are removed, the result of \texttt{l3tl\_range\_unbraced:nnn} may have a different number of items as for \texttt{l3tl\_range:nnn} or \texttt{l3tl\_range\_braced:nnn}. In cases where preserving spaces is important, consider the slower function \texttt{l3tl\_range:nnn}.

\textbf{\TeXhackers note :} The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp_not:n}), which means that the \langle item \rangle does not expand further when appearing in an x-type argument expansion.
\tl_build_put_left:Nn
\tl_build_put_left:Nx
\tl_build_gput_left:Nn
\tl_build_gput_left:Nx
\tl_build_put_right:Nn
\tl_build_put_right:Nx
\tl_build_gput_right:Nn
\tl_build_gput_right:Nx

Rev: 2018-04-01

\tl_build_get:N\tl_build_get:NN

Rev: 2018-04-01

\tl_build_end:N
\tl_build_end:N
\tl_build_gend:N

Rev: 2018-04-01

\c_catcode_active_space_tl
\char_to_utfviii_bytes:n
\char_to_nfd:N

Rev: 2017-08-07

Rev: 2020-01-09

Rev: 2020-01-02

\tl_build_put_left:Nn (\tl var) \{\{\textit{tokens}\}\}
\tl_build_put_right:Nn (\tl var) \{\{\textit{tokens}\}\}

Adds \textit{\{tokens\}} to the left or right side of the current contents of \langle \tl var \rangle. The \langle \tl var \rangle must have been set up with \tl_build_begin:N or \tl_build_gbegin:N. The \textit{put} and \textit{gput} functions must be used for local and global \langle \tl var \rangle respectively. The \textit{right} functions are about twice faster than the \textit{left} functions.

\tl_build_get:N (\tl var_1) (\tl var_2)

Stores the contents of the \langle \tl var_1 \rangle in the \langle \tl var_2 \rangle. The \langle \tl var_1 \rangle must have been set up with \tl_build_begin:N or \tl_build_gbegin:N. The \langle \tl var_2 \rangle is a “normal” token list variable, assigned locally using \tl_set:Nn.

\tl_build_end:N (\tl var)

\tl_build_gend:N

Rev: 2018-04-01

\tl_build_get:N (\tl var_1) (\tl var_2)

Stores the contents of the \langle \tl var_1 \rangle in the \langle \tl var_2 \rangle. The \langle \tl var_1 \rangle must have been set up with \tl_build_begin:N or \tl_build_gbegin:N. The \langle \tl var_2 \rangle is a “normal” token list variable, assigned locally using \tl_set:Nn.

\tl_build_end:N (\tl var)

\tl_build_gend:N

Rev: 2018-04-01

\char_to_utfviii_bytes:n \{\textit{codepoint}\}

Converts the (Unicode) \langle \textit{codepoint} \rangle to UTF-8 bytes. The expansion of this function comprises four brace groups, each of which will contain a hexadecimal value: the appropriate byte. As UTF-8 is a variable-length, one or more of the groups may be empty: the bytes read in the logical order, such that a two-byte codepoint will have groups \#1 and \#2 filled and \#3 and \#4 empty.

\char_to_nfd:N \{\textit{char}\}

Converts the \langle \textit{char} \rangle to the Unicode Normalization Form Canonical Decomposition. The category code of the generated character is the same as the \langle \textit{char} \rangle. With 8-bit engines, no change is made to the character.

14 Additions to \texttt{l3token}

Token list containing one character with category code 13, (“active”), and character code 32 (space).

\char_to_utfviii_bytes:n \{\textit{codepoint}\}

Converts the (Unicode) \langle \textit{codepoint} \rangle to UTF-8 bytes. The expansion of this function comprises four brace groups, each of which will contain a hexadecimal value: the appropriate byte. As UTF-8 is a variable-length, one or more of the groups may be empty: the bytes read in the logical order, such that a two-byte codepoint will have groups \#1 and \#2 filled and \#3 and \#4 empty.

\char_to_nfd:N \{\textit{char}\}

Converts the \langle \textit{char} \rangle to the Unicode Normalization Form Canonical Decomposition. The category code of the generated character is the same as the \langle \textit{char} \rangle. With 8-bit engines, no change is made to the character.
Collects and removes tokens from the input stream until finding a token that does not match the \( \text{⟨test token⟩} \) (as defined by the test \texttt{\token_if_eq_catcode:NNTF} or \texttt{\token_if_eq_charcode:NNTF} or \texttt{\token_if_eq_meaning:NNTF}). The collected tokens are passed to the \( \text{⟨inline code⟩} \) as \#1. When begin-group or end-group tokens (usually \{ or \}) are collected they are replaced by implicit \texttt{\c_group_begin_token} and \texttt{\c_group_end_token}, and when spaces (including \texttt{\c_space_token}) are collected they are replaced by explicit spaces.

For example the following code prints “Hello” to the terminal and leave “, world!” in the input stream.

\[
\begin{align*}
\texttt{\peek_catcode_collect_inline:Nn} & \quad \texttt{\peek_catcode_collect_inline:Nn} \text{ ⟨test token⟩} \{(\text{inline code})}\quad \texttt{\peek_catcode_collect_inline:Nn} \\
\texttt{\peek_charcode_collect_inline:Nn} & \quad \texttt{\peek_charcode_collect_inline:Nn} \text{ ⟨test token⟩} \{(\text{inline code})}\quad \texttt{\peek_charcode_collect_inline:Nn} \\
\texttt{\peek_meaning_collect_inline:Nn} & \quad \texttt{\peek_meaning_collect_inline:Nn} \text{ ⟨test token⟩} \{(\text{inline code})}\quad \texttt{\peek_meaning_collect_inline:Nn} \\
\end{align*}
\]

\texttt{\peek_meaning_collect_inline:Nn} \texttt{\c_space_token} \{ \texttt{\peek_charcode:NTF} \texttt{*} \{ \texttt{star} \} \{ \texttt{no-star \#1} \} \}

Removes explicit and implicit space tokens (category code 10 and character code 32) from the input stream, then inserts \( \text{⟨code⟩} \).
## Index

The italic numbers denote the pages where the corresponding entry is described, numbers underlined point to the definition, all others indicate the places where it is used.

<table>
<thead>
<tr>
<th>Symbols</th>
<th>atand</th>
<th>\begin</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>!</td>
<td>211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;&amp;</td>
<td>211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>211, 210, 211</td>
<td>\bool_case_false:n</td>
<td>264</td>
</tr>
<tr>
<td>-</td>
<td>211, 210, 211</td>
<td>\bool_case_false:nTF</td>
<td>264</td>
</tr>
<tr>
<td>/</td>
<td>211</td>
<td>\bool_case_true:nTF</td>
<td>264</td>
</tr>
<tr>
<td>::</td>
<td>\texttt{\begin}:</td>
<td>\bool_case_true:nTF</td>
<td>264</td>
</tr>
<tr>
<td>::V</td>
<td>37</td>
<td>\bool_const:N</td>
<td>107</td>
</tr>
<tr>
<td>::V_unbraced</td>
<td>37</td>
<td>\bool_do_until:Nn</td>
<td>110</td>
</tr>
<tr>
<td>::c</td>
<td>37</td>
<td>\bool_do_until:nn</td>
<td>111</td>
</tr>
<tr>
<td>::e</td>
<td>37</td>
<td>\bool_do_while:Nn</td>
<td>110</td>
</tr>
<tr>
<td>::e_unbraced</td>
<td>37</td>
<td>\bool_do_while:nn</td>
<td>111</td>
</tr>
<tr>
<td>::error</td>
<td>263</td>
<td>\bool_gset:N</td>
<td>185</td>
</tr>
<tr>
<td>::f</td>
<td>37</td>
<td>\bool_gset:Nn</td>
<td>107</td>
</tr>
<tr>
<td>::f_unbraced</td>
<td>37</td>
<td>\bool_gset_eq:NN</td>
<td>107</td>
</tr>
<tr>
<td>::n</td>
<td>37</td>
<td>\bool_gset_inverse:N</td>
<td>185</td>
</tr>
<tr>
<td>::n_unbraced</td>
<td>37</td>
<td>\bool_gset_inverse:N</td>
<td>264</td>
</tr>
<tr>
<td>::o</td>
<td>37</td>
<td>\bool_gset_true:N</td>
<td>107</td>
</tr>
<tr>
<td>::o_unbraced</td>
<td>37</td>
<td>\bool_if:NTF</td>
<td>107</td>
</tr>
<tr>
<td>::p</td>
<td>37</td>
<td>\bool_if:nTF</td>
<td>107</td>
</tr>
<tr>
<td>::p_v</td>
<td>37</td>
<td>\bool_if:nTF 107, 109, 111, 111, 111</td>
<td>110</td>
</tr>
<tr>
<td>::p_v_unbraced</td>
<td>37</td>
<td>\bool_if_exist:NTF</td>
<td>107</td>
</tr>
<tr>
<td>::p_x</td>
<td>37</td>
<td>\bool_if_exist_p:N</td>
<td>108</td>
</tr>
<tr>
<td>::p_x_unbraced</td>
<td>37</td>
<td>\bool_if_p:N</td>
<td>110</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_if_p:n</td>
<td>109</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_lazy_all:nTF</td>
<td>109, 109, 109</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_lazy_all_p:n</td>
<td>109</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_lazy_and:nTF</td>
<td>109, 109, 109</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_lazy_and_p:n</td>
<td>109, 109, 109</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_lazy_or:nTF 109, 109, 110, 110</td>
<td>110</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_lazy_or_p:n</td>
<td>110</td>
</tr>
<tr>
<td>:&lt; :&gt;</td>
<td>210</td>
<td>\bool_log:N</td>
<td>107</td>
</tr>
<tr>
<td>abs</td>
<td>211</td>
<td>\bool_log:n</td>
<td>108</td>
</tr>
<tr>
<td>acos</td>
<td>213</td>
<td>\bool_new:N</td>
<td>106</td>
</tr>
<tr>
<td>acosd</td>
<td>213</td>
<td>\bool_not:p:n</td>
<td>110</td>
</tr>
<tr>
<td>acot</td>
<td>214</td>
<td>\bool_set:N</td>
<td>185</td>
</tr>
<tr>
<td>acotd</td>
<td>214</td>
<td>\bool_set:Nn</td>
<td>107</td>
</tr>
<tr>
<td>acsc</td>
<td>213</td>
<td>\bool_set_eq:NN</td>
<td>107</td>
</tr>
<tr>
<td>acscd</td>
<td>213</td>
<td>\bool_set_eq:N</td>
<td>107</td>
</tr>
<tr>
<td>asec</td>
<td>213</td>
<td>\bool_set_inverse:N</td>
<td>185</td>
</tr>
<tr>
<td>asecd</td>
<td>213</td>
<td>\bool_set_inverse:N</td>
<td>264</td>
</tr>
<tr>
<td>asind</td>
<td>213</td>
<td>\bool_set_true:N</td>
<td>107</td>
</tr>
<tr>
<td>asind</td>
<td>213</td>
<td>\bool_set_true:N</td>
<td>107</td>
</tr>
<tr>
<td>atan</td>
<td>214</td>
<td>\bool_show:N</td>
<td>107</td>
</tr>
<tr>
<td>atan</td>
<td>214</td>
<td>\bool_show:n</td>
<td>107</td>
</tr>
</tbody>
</table>

271
\ExplFileDescription ...... 7 \flag_if_exist:n ............ 103
\ExplFileName ............... 7 \flag_if_exist:nTF ............ 103
\ExplFileVersion ............ 7 \flag_if_exist:pn .......... 103
\ExplSyntaxOff ............ 4, 7, 118 \flag_if_raised:n ............ 103
\ExplSyntaxOn ............ 4, 7, 118 \flag_if_raised:nTF .......... 103

\flag_new:n ............ 102, 102 \flag_raise:n ............ 103
\flag_raise_if_clear:n .... 262 \flag_show:n ............ 102

F
\vspace{1em}
\begin{verbatim}
\textbf{F i commands:}
\end{verbatim}

\begin{verbatim}
\textbf{file commands:}
\end{verbatim}

\begin{verbatim}
\textbf{fp commands:}
\end{verbatim}

\begin{verbatim}
\textbf{flag commands:}
\end{verbatim}

\begin{verbatim}
\end{verbatim}
\fp_new:N 199, 200
\fp_set:N 186
\fp_set:Nn 200
\fp_set_eq:NW 200
\fp_show:N 207
\fp_show:n 207
\fp_sign:n 201
\fp_show:n 201
\fp_set_eq:NN 200
\fp_set:Nn 200
\fp_new:N 200
\fparray_new:Nn 219
\fparray_item_to_tl:Nn 219
\fparray_gzero:N 216
\fparray_count:N 219
\__fp_basics_pack_weird_low:NNNNw 219
\c_zero_fp 219
\l_tmpb_fp 219
\g_tmpb_fp 219
\l_tmpa_fp 219
\c_pi_fp 219
\c_nan_fp 219
\c_one_fp 219
\c_pi_fp 219
\g_tapa_fp 219
\l_tapa_fp 219
\g_tampb_fp 219
\l_tampb_fp 219
\c_zero_fp 219
\c_inf_fp 206, 215
\c_nan_fp 215
\c_one_fp 205
\c_pi_fp 206, 215
\g_tapa_fp 206
\l_tapa_fp 206
\g_tampb_fp 206
\l_tampb_fp 206
\c_zero_fp 205
\_fp_basics_pack_weird_high:NNNNNNNN
\_fp_basics_pack_weird_low:NNNN
\fparray_count:N 219, 219, 219
\fparray_gset:Nnn 219
\fparray_gzero:N 219
\fparray_item:Nn 219
\fparray_item_to_tl:Nn 219
\fparray_new:Nn 219
\GetIdInfo 7
\prg_set_protected_conditional:Npn
\c_empty_prop
\prop_clear:N
\prop_clear_new:N
\prop_const_from_keyval:Nn
\prop_count:N
\prop_gclear:N
\prop_gclear_new:N
\prop_get:Nn
\prop_get:NnN
\prop_gpop:NnN
\prop_gpop:NnTF
\prop_gput:N
\prop_gput:Nnn
\prop_gput_if_new:Nnn
\prop_gremove:Nn
\prop_get_eq:NN
\prop_get_from_keyval:Nn
\prop_if_empty:NTF
\prop_if_empty:p:N
\prop_if_exist:NTF
\prop_if_exist:p:N
\prop_if_in:NnTF
\prop_if_in:p:Nn
\prop_item:Nn
\prop_log:N
\prop_map_break:
\prop_map_break:n
\prop_map_function:NN
\prop_map_inline:Nn
\prop_map_tokens:Nn
\prop_new:N
\prop_pop:NnN
\prop_pop:NnTF
\prop_put:N
\prop_put:Nnn
\prop_put_if_new:Nnn
\prop_rand_key_value:N
\prop_remove:Nn
\prop_set_eq:NN
\prop_set_from_keyval:Nn
\prop_show:N
\g_tempa_prop
\g_tempb_prop
\g_tempc_prop
\l_tempa_prop
\l_tempb_prop
\l_tempc_prop
\ProvidesExplClass
\ProvidesExplFile
\ProvidesExplPackage
\prop
\prop commands:
\q_mark
\q_nil
\q_no_value
\q_if_nil:NTF
\q_if_nil:nTF
\q_if_nil:p:N
\q_if_nil:p:n
\q_if_no_value:NTF
\q_if_no_value:nTF
\q_if_no_value:p:N
\q_if_no_value:p:n
\q_if_recursion_tail_-
\break:N
\q_if_recursion_tail_-
\break:N
\q_if_recursion_tail_stop:N
\q_if_recursion_tail_stop:n
\do:Nn
\q_if_recursion_tail_stop_-
\do:nn
\q_recursion_stop
\q_recursion_tail
\s_stop
\rand
\randint
\regex commands:
\c_foo_regex
\regex_gset:Nn
\regex_extract_all:nnN
\regex_const:Nn
\regex_(g)set:Nn
\regex_count:Nn
\regex_count:nnN
\regex_extract_all:NnN
\regex_extract_all:nnN
\regex_extract_all:NnTF
\regex_extract_all:nnTF
\regex_extract_all_once:Nn
\regex_extract_once:nnN
\regex_extract_once:NnTF
\regex_extract_once:nnTF
\ProvidesExplClass
\ProvidesExplFile
\ProvidesExplPackage
\rand
\randint
\regex commands:
sys commands:
\sys_gset_rand_seed:n ........................... 58
\sys_get_shell:nnTF ............................ 58
\sys_get_shell:nnN(TF) .......................... 58
\sys_item:nn ................................. 62
\sys_item:n ................................. 62
\sys_item_ignore_spaces:nn .......................... 62
\sys_log:N .............................. 65
\sys_log:n .............................. 65
\sys_lowercase:n .............................. 64, 257
\sys_map_break: ............................... 60
\sys_map_break:n ............................ 60, 61
\sys_map_function:NN .......................... 59, 59, 60, 60
\sys_map_function:nN ........................... 59, 59
\sys_map_inline:Nnn ............................ 60
\sys_map_variable:NNn .......................... 60
\sys_new:N ............................. 55, 56
\sys_new:Nn .............................. 56
\sys_new:Nnn .............................. 57
\sys_range:Nnn ............................ 63
\sys_range:nnnn ............................ 63, 165
\sys_range_ignore_spaces:nnnn .......................... 63
\sys_remove_all:Nn ............................. 57, 57
\sys_remove_once:Nn ............................. 57
\sys_replace_all:Nnn .......................... 57
\sys_replace_once:Nnnn ........................ 57
\sys_set:Nn ............................... 56, 57
\sys_set_convert:Nnmnm .......................... 67, 69, 69
\sys_set_convert:NnnmTF .......................... 69
\sys_set_eq:NN ............................... 56
\sys_show:N .............................. 65
\sys_show:n .............................. 65
\sys_stat:N .............................. 62
\sys_stat:n .............................. 62
\sys_tilde:nnn .............................. 62
\sys_tilde:nnnn ............................. 63
\sys_use:N ............................. 61
\c_tilde_str .............................. 66
\g_tilde_str .............................. 66
\l_tilde_str .............................. 57, 66
\g_tamp_str .............................. 66
\l_tamp_str .............................. 66
\c_underscore_str ............................ 66
tan .............................................. 213
tand .............................................. 213
Tex and \TeX\texttt{2e} commands:
\TeX  ........................................ 19
\TeX  ........................................ 19
\filelist  .................................... 167
\firstofone ................................ 19
\firstoftwo  ................................ 19
\gobble ................................... 20
\gobbletwo ................................... 20
\secondoftwo ................................ 19
\box ........................................ 242
\char ....................................... 141

286
U

undefined commands:
  .undefine: .......................... 188

use commands:
  \use:N .......................... 16, 102
  \use:n ...................... 19, 20, 20, 38, 140
  \use:nn ......................... 19
  \use:nnn ....................... 19
  \use:i:nn ...................... 19
  \use:i:nmm .................... 19
  \use:i:nnn .................... 19
  \use:i:nnnn ................... 19
  \use_i:nn ...................... 19
  \use_i:nmm .................... 19
  \use_i:nnn .................... 19
  \use_i:nnnn ................... 19
  \use_i_delimit_by_q_nil:nw .... 21
  \use_i_delimit_by_q_recursion_-  
    stop:nw ....................... 21
  \use_i_delimit_by_q_recursion_-  
    stop:w ...................... 72, 72
  \use_i_delimit_by_q_stop:nw .... 21
  \use_i_delimit_by_q_stop:w ..... 21
  \use_i_delimit_by_s_stop:w .... 74, 74

V

value commands:
  .value_forbidden:n ............. 188
  .value_required:n ............ 188

vbox commands:
  \vbox:n .................. 236, 236, 240
  \vbox_gset:Nn ............ 240
  \vbox_gset:Nw ............ 241
  \vbox_gset_end: .......... 241
  \vbox_gset_split_to_ht:NNn .... 241
  \vbox_gset_to_ht:Nnn ..... 240
  \vbox_gset_to_ht:Nnw ..... 241
  \vbox_gset_top:Nn .. 240
  \vbox_set:Nn .......... 240, 241
  \vbox_set:Nw .......... 241
  \vbox_set_end: .......... 241, 241
  \vbox_set_split_to_ht:NNn .... 241
  \vbox_set_to_ht:Nnn ..... 240, 241
  \vbox_set_to_ht:Nnw ..... 241
  \vbox_set_top:Nn ...... 240
  \vbox_to_ht:nn ........ 240
  \vbox_to_zero:n .......... 240
  \vbox_top:n .......... 240
  \vbox_unpack:N .......... 241
  \vbox_unpack_drop:N ...... 242

vcoffin commands:
  \vcoffin_gset:Nnn .......... 247
  \vcoffin_gset:Nww .......... 247
  \vcoffin_gset_end: ...... 247
  \vcoffin_set:Nn .......... 247
  \vcoffin_set:Nww .......... 247
  \vcoffin_set_end: ...... 247