The \LaTeX3 Interfaces

The \LaTeX Project*

Released 2021-06-01

Abstract

This is the reference documentation for the expl3 programming environment. The 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 \eps-\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 expl3 modules define an independent low-level \LaTeX3 programming language.

The expl3 modules are designed to be loaded on top of \LaTeX2e. With an up-to-date \LaTeX2e kernel, this material is loaded as part of the format. The fundamental programming code can also be loaded with other \TeX formats, subject to restrictions on the full range of functionality.

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

I Introduction 1

1 Introduction to expl3 and this document 2
  1.1 Naming functions and variables .................................. 2
  1.1.1 Terminological inexactitude .................................. 5
  1.2 Documentation conventions ....................................... 5
  1.3 Formal language conventions which apply generally ............. 7
  1.4 \TeX concepts not supported by \LaTeX3 .......................... 7

II Bootstrapping 8

2 The l3bootstrap package: Bootstrap code 9
  2.1 Using the \LaTeX3 modules ....................................... 9

3 The l3names package: Namespace for primitives 11
  3.1 Setting up the \LaTeX3 programming language ..................... 11

III Programming Flow 12

4 The l3basics package: Basic definitions 13
  4.1 No operation functions .......................................... 13
  4.2 Grouping material ............................................. 13
  4.3 Control sequences and functions ................................ 14
    4.3.1 Defining functions ....................................... 14
    4.3.2 Defining new functions using parameter text .......... 15
    4.3.3 Defining new functions using the signature ........... 17
    4.3.4 Copying control sequences ............................... 19
    4.3.5 Deleting control sequences ............................... 20
    4.3.6 Showing control sequences ............................... 20
    4.3.7 Converting to and from control sequences ............. 20
  4.4 Analysing control sequences ................................... 22
  4.5 Using or removing tokens and arguments ........................ 23
    4.5.1 Selecting tokens from delimited arguments ............ 25
  4.6 Predicates and conditionals ................................... 25
    4.6.1 Tests on control sequences ............................... 27
    4.6.2 Primitive conditionals ................................... 27
  4.7 Starting a paragraph ........................................... 28
  4.8 Debugging support ............................................. 29
5 The \texttt{l3expan} package: Argument expansion
\begin{itemize}
\item 5.1 Defining new variants .............................................. 30
\item 5.2 Methods for defining variants .................................. 31
\item 5.3 Introducing the variants ....................................... 32
\item 5.4 Manipulating the first argument ............................... 34
\item 5.5 Manipulating two arguments ................................. 36
\item 5.6 Manipulating three arguments ............................... 37
\item 5.7 Unbraced expansion ............................................ 38
\item 5.8 Preventing expansion .......................................... 38
\item 5.9 Controlled expansion ......................................... 40
\item 5.10 Internal functions ........................................... 42
\end{itemize}

6 The \texttt{l3sort} package: Sorting functions ..................... 43
\begin{itemize}
\item 6.1 Controlling sorting ............................................... 43
\end{itemize}

7 The \texttt{l3tl-analysis} package: Analysing token lists .......... 45

8 The \texttt{l3regex} package: Regular expressions in \TeX{} ......... 46
\begin{itemize}
\item 8.1 Syntax of regular expressions ............................... 47
\item 8.1.1 Regex examples ........................................... 47
\item 8.1.2 Characters in regular expressions ....................... 48
\item 8.1.3 Characters classes ..................................... 48
\item 8.1.4 Structure: alternatives, groups, repetitions ............ 49
\item 8.1.5 Matching exact tokens ................................. 50
\item 8.1.6 Miscellaneous ........................................... 52
\item 8.2 Syntax of the replacement text ........................... 52
\item 8.3 Pre-compiling regular expressions ......................... 54
\item 8.4 Matching .................................................. 54
\item 8.5 Submatch extraction ...................................... 55
\item 8.6 Replacement ............................................... 56
\item 8.7 Constants and variables .................................. 57
\item 8.8 Bugs, misfeatures, future work, and other possibilities . 57
\end{itemize}

9 The \texttt{l3prg} package: Control structures ...................... 60
\begin{itemize}
\item 9.1 Defining a set of conditional functions ..................... 60
\item 9.2 The boolean data type ..................................... 62
\item 9.2.1 Scratch booleans .................................... 64
\item 9.3 Boolean expressions ....................................... 64
\item 9.4 Logical loops ............................................. 66
\item 9.5 Producing multiple copies .................................. 67
\item 9.6 Detecting \TeX{}'s mode .................................. 68
\item 9.7 Primitive conditionals ................................... 68
\item 9.8 Nestable recursions and mappings ......................... 68
\item 9.8.1 Simple mappings .................................... 69
\item 9.9 Internal programming functions ........................... 69
\end{itemize}
10 The \texttt{l3sys} package: System/runtime functions 70
10.1 The name of the job ........................................... 70
10.2 Date and time .................................................. 70
10.3 Engine .......................................................... 71
10.4 Output format .................................................. 71
10.5 Platform ......................................................... 72
10.6 Random numbers ............................................... 72
10.7 Access to the shell ............................................ 72
10.8 Loading configuration data ................................. 73
10.8.1 Final settings ............................................... 74

11 The \texttt{l3msg} package: Messages 75
11.1 Creating new messages ...................................... 75
11.2 Customizable information for message modules .......... 76
11.3 Contextual information for messages .................... 76
11.4 Issuing messages .............................................. 78
11.4.1 Messages for showing material ....................... 81
11.4.2 Expandable error messages ........................... 81
11.5 Redirecting messages ....................................... 81

12 The \texttt{l3file} package: File and I/O operations 83
12.1 Input–output stream management ......................... 83
12.1.1 Reading from files ..................................... 85
12.1.2 Writing to files ......................................... 88
12.1.3 Wrapping lines in output .............................. 90
12.1.4 Constant input–output streams, and variables .... 91
12.1.5 Primitive conditionals ................................ 91
12.2 File operation functions ................................... 91

13 The \texttt{l3luatex} package: Lua\TeX{}-specific functions 96
13.1 Breaking out to Lua ......................................... 96
13.2 Lua interfaces ............................................... 97

14 The \texttt{l3legacy} package: Interfaces to legacy concepts 99

IV Data types 100
15 The \texttt{l3tl} package: Token lists 101
15.1 Creating and initialising token list variables .......... 101
15.2 Adding data to token list variables .................... 102
15.3 Modifying token list variables .......................... 103
15.4 Reassigning token list category codes ................ 103
15.5 Token list conditionals .................................. 105
15.6 Mapping over token lists .................................. 107
15.7 Using token lists .......................................... 109
15.8 Working with the content of token lists ............... 109
15.9 The first token from a token list ...................... 111
15.10 Using a single item ...................................... 114
15.11 Viewing token lists ...................................... 116
16 The l3str package: Strings
  16.1 Building strings ........................................ 119
  16.2 Adding data to string variables .......................... 120
  16.3 Modifying string variables ................................ 121
  16.4 String conditionals ....................................... 122
  16.5 Mapping over strings ...................................... 123
  16.6 Working with the content of strings ..................... 124
  16.7 String manipulation ...................................... 128
  16.8 Viewing strings .......................................... 129
  16.9 Constant token lists ..................................... 130
  16.10 Scratch strings ......................................... 130

17 The l3str-convert package: string encoding conversions 131
  17.1 Encoding and escaping schemes ......................... 131
  17.2 Conversion functions .................................... 133
  17.3 Conversion by expansion (for PDF contexts) .......... 133
  17.4 Possibilities, and things to do .......................... 133

18 The l3quark package: Quarks 135
  18.1 Quarks .................................................. 135
  18.2 Defining quarks ......................................... 136
  18.3 Quark tests .............................................. 136
  18.4 Recursion ................................................ 137
    18.4.1 An example of recursion with quarks ............. 138
  18.5 Scan marks .............................................. 139

19 The l3seq package: Sequences and stacks 140
  19.1 Creating and initialising sequences .................... 140
  19.2 Appending data to sequences ............................ 142
  19.3 Recovering items from sequences ....................... 142
  19.4 Recovering values from sequences with branching ..... 143
  19.5 Modifying sequences .................................... 145
  19.6 Sequence conditionals .................................. 145
  19.7 Mapping over sequences .................................. 146
  19.8 Using the content of sequences directly ............... 148
  19.9 Sequences as stacks ..................................... 149
  19.10 Sequences as sets ....................................... 150
  19.11 Constant and scratch sequences ....................... 151
  19.12 Viewing sequences ..................................... 152
20 The l3int package: Integers 153
20.1 Integer expressions ................................. 154
20.2 Creating and initialising integers .................. 155
20.3 Setting and incrementing integers ................. 156
20.4 Using integers ...................................... 157
20.5 Integer expression conditionals .................. 157
20.6 Integer expression loops ........................... 159
20.7 Integer step functions .............................. 161
20.8 Formatting integers .................................. 162
20.9 Converting from other formats to integers ...... 163
20.10 Random integers .................................... 164
20.11 Viewing integers .................................... 165
20.12 Constant integers ................................... 165
20.13 Scratch integers .................................... 165
20.14 Direct number expansion ........................... 166
20.15 Primitive conditionals .............................. 166

21 The l3flag package: Expandable flags 168
21.1 Setting up flags .................................... 168
21.2 Expandable flag commands ......................... 169

22 The l3clist package: Comma separated lists 170
22.1 Creating and initialising comma lists ............ 171
22.2 Adding data to comma lists ......................... 172
22.3 Modifying comma lists .............................. 173
22.4 Comma list conditionals ............................ 174
22.5 Mapping over comma lists ........................... 174
22.6 Using the content of comma lists directly ...... 176
22.7 Comma lists as stacks .............................. 177
22.8 Using a single item .................................. 178
22.9 Viewing comma lists .................................. 179
22.10 Constant and scratch comma lists ............... 179

23 The l3token package: Token manipulation 180
23.1 Creating character tokens ......................... 181
23.2 Manipulating and interrogating character tokens 182
23.3 Generic tokens ...................................... 185
23.4 Converting tokens ................................... 186
23.5 Token conditionals ................................... 186
23.6 Peeking ahead at the next token ................. 190
23.7 Description of all possible tokens ............... 195
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>The l3prop package: Property lists</td>
<td>198</td>
</tr>
<tr>
<td>24.1</td>
<td>Creating and initialising property lists</td>
<td>198</td>
</tr>
<tr>
<td>24.2</td>
<td>Adding and updating property list entries</td>
<td>199</td>
</tr>
<tr>
<td>24.3</td>
<td>Recovering values from property lists</td>
<td>200</td>
</tr>
<tr>
<td>24.4</td>
<td>Modifying property lists</td>
<td>201</td>
</tr>
<tr>
<td>24.5</td>
<td>Property list conditionals</td>
<td>201</td>
</tr>
<tr>
<td>24.6</td>
<td>Recovering values from property lists with branching</td>
<td>202</td>
</tr>
<tr>
<td>24.7</td>
<td>Mapping over property lists</td>
<td>202</td>
</tr>
<tr>
<td>24.8</td>
<td>Viewing property lists</td>
<td>204</td>
</tr>
<tr>
<td>24.9</td>
<td>Scratch property lists</td>
<td>204</td>
</tr>
<tr>
<td>24.10</td>
<td>Constants</td>
<td>205</td>
</tr>
<tr>
<td>25</td>
<td>The l3skip package: Dimensions and skips</td>
<td>206</td>
</tr>
<tr>
<td>25.1</td>
<td>Creating and initialising dim variables</td>
<td>206</td>
</tr>
<tr>
<td>25.2</td>
<td>Setting dim variables</td>
<td>207</td>
</tr>
<tr>
<td>25.3</td>
<td>Utilities for dimension calculations</td>
<td>207</td>
</tr>
<tr>
<td>25.4</td>
<td>Dimension expression conditionals</td>
<td>208</td>
</tr>
<tr>
<td>25.5</td>
<td>Dimension expression loops</td>
<td>210</td>
</tr>
<tr>
<td>25.6</td>
<td>Dimension step functions</td>
<td>211</td>
</tr>
<tr>
<td>25.7</td>
<td>Using dim expressions and variables</td>
<td>212</td>
</tr>
<tr>
<td>25.8</td>
<td>Viewing dim variables</td>
<td>213</td>
</tr>
<tr>
<td>25.9</td>
<td>Constant dimensions</td>
<td>214</td>
</tr>
<tr>
<td>25.10</td>
<td>Scratch dimensions</td>
<td>214</td>
</tr>
<tr>
<td>25.11</td>
<td>Creating and initialising skip variables</td>
<td>214</td>
</tr>
<tr>
<td>25.12</td>
<td>Setting skip variables</td>
<td>215</td>
</tr>
<tr>
<td>25.13</td>
<td>Skip expression conditionals</td>
<td>216</td>
</tr>
<tr>
<td>25.14</td>
<td>Using skip expressions and variables</td>
<td>216</td>
</tr>
<tr>
<td>25.15</td>
<td>Viewing skip variables</td>
<td>216</td>
</tr>
<tr>
<td>25.16</td>
<td>Constant skips</td>
<td>217</td>
</tr>
<tr>
<td>25.17</td>
<td>Scratch skips</td>
<td>217</td>
</tr>
<tr>
<td>25.18</td>
<td>Inserting skips into the output</td>
<td>217</td>
</tr>
<tr>
<td>25.19</td>
<td>Creating and initialising muskip variables</td>
<td>218</td>
</tr>
<tr>
<td>25.20</td>
<td>Setting muskip variables</td>
<td>218</td>
</tr>
<tr>
<td>25.21</td>
<td>Using muskip expressions and variables</td>
<td>219</td>
</tr>
<tr>
<td>25.22</td>
<td>Viewing muskip variables</td>
<td>219</td>
</tr>
<tr>
<td>25.23</td>
<td>Constant muskips</td>
<td>220</td>
</tr>
<tr>
<td>25.24</td>
<td>Scratch muskips</td>
<td>220</td>
</tr>
<tr>
<td>25.25</td>
<td>Primitive conditional</td>
<td>220</td>
</tr>
<tr>
<td>26</td>
<td>The l3keys package: Key–value interfaces</td>
<td>221</td>
</tr>
<tr>
<td>26.1</td>
<td>Creating keys</td>
<td>222</td>
</tr>
<tr>
<td>26.2</td>
<td>Sub-dividing keys</td>
<td>226</td>
</tr>
<tr>
<td>26.3</td>
<td>Choice and multiple choice keys</td>
<td>227</td>
</tr>
<tr>
<td>26.4</td>
<td>Setting keys</td>
<td>229</td>
</tr>
<tr>
<td>26.5</td>
<td>Handling of unknown keys</td>
<td>230</td>
</tr>
<tr>
<td>26.6</td>
<td>Selective key setting</td>
<td>231</td>
</tr>
<tr>
<td>26.7</td>
<td>Utility functions for keys</td>
<td>232</td>
</tr>
<tr>
<td>26.8</td>
<td>Low-level interface for parsing key–val lists</td>
<td>232</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
<td>Pages</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>27</td>
<td>The \texttt{l3intarray} package: fast global integer arrays</td>
<td>235</td>
</tr>
<tr>
<td>27.1</td>
<td>\texttt{l3intarray} documentation</td>
<td>235</td>
</tr>
<tr>
<td>27.1.1</td>
<td>Implementation notes</td>
<td>236</td>
</tr>
<tr>
<td>28</td>
<td>The \texttt{l3fp} package: Floating points</td>
<td>237</td>
</tr>
<tr>
<td>28.1</td>
<td>Creating and initialising floating point variables</td>
<td>239</td>
</tr>
<tr>
<td>28.2</td>
<td>Setting floating point variables</td>
<td>239</td>
</tr>
<tr>
<td>28.3</td>
<td>Using floating points</td>
<td>240</td>
</tr>
<tr>
<td>28.4</td>
<td>Floating point conditionals</td>
<td>241</td>
</tr>
<tr>
<td>28.5</td>
<td>Floating point expression loops</td>
<td>243</td>
</tr>
<tr>
<td>28.6</td>
<td>Some useful constants, and scratch variables</td>
<td>245</td>
</tr>
<tr>
<td>28.7</td>
<td>Floating point exceptions</td>
<td>246</td>
</tr>
<tr>
<td>28.8</td>
<td>Viewing floating points</td>
<td>247</td>
</tr>
<tr>
<td>28.9</td>
<td>Floating point expressions</td>
<td>248</td>
</tr>
<tr>
<td>28.9.1</td>
<td>Input of floating point numbers</td>
<td>248</td>
</tr>
<tr>
<td>28.9.2</td>
<td>Precedence of operators</td>
<td>249</td>
</tr>
<tr>
<td>28.9.3</td>
<td>Operations</td>
<td>249</td>
</tr>
<tr>
<td>28.10</td>
<td>Disclaimer and roadmap</td>
<td>256</td>
</tr>
<tr>
<td>29</td>
<td>The \texttt{l3fparray} package: fast global floating point arrays</td>
<td>259</td>
</tr>
<tr>
<td>29.1</td>
<td>\texttt{l3fparray} documentation</td>
<td>259</td>
</tr>
<tr>
<td>30</td>
<td>The \texttt{l3ctab} package: Category code tables</td>
<td>260</td>
</tr>
<tr>
<td>30.1</td>
<td>Creating and initialising category code tables</td>
<td>260</td>
</tr>
<tr>
<td>30.2</td>
<td>Using category code tables</td>
<td>260</td>
</tr>
<tr>
<td>30.3</td>
<td>Category code table conditionals</td>
<td>261</td>
</tr>
<tr>
<td>30.4</td>
<td>Constant category code tables</td>
<td>261</td>
</tr>
<tr>
<td>V</td>
<td>Text manipulation</td>
<td>262</td>
</tr>
<tr>
<td>31</td>
<td>The \texttt{l3unicode} package: Unicode support functions</td>
<td>263</td>
</tr>
<tr>
<td>32</td>
<td>The \texttt{l3text} package: text processing</td>
<td>264</td>
</tr>
<tr>
<td>32.1</td>
<td>Expanding text</td>
<td>264</td>
</tr>
<tr>
<td>32.2</td>
<td>Case changing</td>
<td>265</td>
</tr>
<tr>
<td>32.3</td>
<td>Removing formatting from text</td>
<td>266</td>
</tr>
<tr>
<td>32.4</td>
<td>Control variables</td>
<td>266</td>
</tr>
<tr>
<td>VI</td>
<td>Typesetting</td>
<td>268</td>
</tr>
</tbody>
</table>
# 33 The \texttt{l3box} package: Boxes

- 33.1 Creating and initialising boxes 
- 33.2 Using boxes
- 33.3 Measuring and setting box dimensions
- 33.4 Box conditionals
- 33.5 The last box inserted
- 33.6 Constant boxes
- 33.7 Scratch boxes
- 33.8 Viewing box contents
- 33.9 Boxes and color
- 33.10 Horizontal mode boxes
- 33.11 Vertical mode boxes
- 33.12 Using boxes efficiently
- 33.13 Affine transformations
- 33.14 Primitive box conditionals

# 34 The \texttt{l3coffins} package: Coffin code layer

- 34.1 Creating and initialising coffins
- 34.2 Setting coffin content and poles
- 34.3 Coffin affine transformations
- 34.4 Joining and using coffins
- 34.5 Measuring coffins
- 34.6 Coffin diagnostics
- 34.7 Constants and variables

# 35 The \texttt{l3color} package: Color support

- 35.1 Color in boxes
- 35.2 Color models
- 35.3 Color expressions
- 35.4 Named colors
- 35.5 Selecting colors
- 35.6 Colors for fills and strokes
- 35.7 Multiple color models
- 35.8 Exporting color specifications
- 35.9 Creating new color models

# 36 The \texttt{l3pdf} package: Core PDF support

- 36.1 Objects
- 36.2 Version
- 36.3 Compression
- 36.4 Destinations

# VII Additions and removals
37 The l3candidates package: Experimental additions to l3kernel 299
  37.1 Important notice ........................................... 299
  37.2 Additions to l3box ........................................... 300
  37.3 Additions to l3expan ......................................... 300
  37.4 Additions to l3fp ............................................ 300
  37.5 Additions to l3file ........................................... 301
  37.6 Additions to l3flag ........................................... 301
  37.7 Additions to l3intarray ...................................... 301
  37.8 Additions to l3msg ............................................ 302
  37.9 Additions to l3prg ............................................ 302
  37.10 Additions to l3prop ........................................... 303
  37.11 Additions to l3seq ........................................... 304
  37.12 Additions to l3sys ........................................... 305
  37.13 Additions to l3tl ............................................ 306
  37.14 Additions to l3token ........................................ 307

Index 309
Part I
Introduction
Chapter 1

Introduction to \texttt{expl3} and this document

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

1.1 Naming functions and variables

\LaTeX3 does not use \texttt{@} 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 \textit{signature} of the function.

Each function name starts with the \textit{module} to which it belongs. Thus apart from a small number of very basic functions, all \texttt{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 \texttt{clist} and begin \texttt{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:

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

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

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

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

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-LuaTeX engines older than 2019.

The f specifier stands for full expansion, and in contrast to 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 ⟨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

\texttt{\tl_set:Nn \l_mya_tl { A }}
\texttt{\tl_set:Nn \l_myb_tl { B }}
\texttt{\tl_set:Nf \l_mya_tl { \l_mya_tl \l_myb_tl }}

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.

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

The letter p indicates TeX parameters. Normally this will be used for delimited functions as expl3 provides better methods for creating simple sequential arguments.

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

D The D stands for Do not use. All of the TeX primitives are initially \texttt{\let} to a D name, and some are then given a second name. These functions have no standardized syntax, they are engine dependent and their name can change without warning, thus their use is strongly discouraged in package code: programmers should instead use the interfaces documented in interface3.pdf1.

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

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

1If a primitive offers a functionality not yet in the kernel, programmers and users are encouraged to write to the \LaTeX-L mailing list (mailto:LATEX-L@listserv.uni-heidelberg.de) describing their use-case and intended behaviour, so that a possible interface can be discussed. Temporarily, while an interface is not provided, programmers may use the procedure described in the l3styleguide.pdf.
c Constant: global parameters whose value should not be changed.

g Parameters whose value should only be set globally.

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\(^2\) name and then a descriptive part. Variables end with a short identifier to show the variable type:

clist Comma separated list.

dim “Rigid” lengths.

fp Floating-point values;

int Integer-valued count register.

muskip “Rubber” lengths for use in mathematics.

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

skip “Rubber” lengths.

str String variables: contain character data.

tl Token list variables: placeholder for a token list.

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

bool Either true or false.

box Box register.

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

flag Integer that can be incremented expandably.

fparray Fixed-size array of floating point values.

intarray Fixed-size array of integers.

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

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

regex Regular expression.

\(^2\)The module names are not used in case of generic scratch registers defined in the data type modules, e.g., the int module contains some scratch variables called \texttt{\_tmpa\_int}, \texttt{\_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{\_int\_tmpa\_int} would be very unreadable.
1.1.1 Terminological inexactitude

A word of warning. In this document, and others referring to the 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. 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 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 TeXbook 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.

1.2 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
\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: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.

\textsuperscript{3}TeXnically, functions with no arguments are \texttt{long} while token list variables are not.
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 \texttt{edef} or \texttt{\expanded}), as well as within an \(f\)-type argument. These fully expandable functions are indicated in the documentation by a star:

\begin{verbatim}
\cs_to_str:N ✯ \cs_to_str:N \langle cs \rangle
\end{verbatim}

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 \langle cs \rangle, shorthand for a \langle control sequence \rangle.

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:

\begin{verbatim}
\seq_map_function:NN ✯ \seq_map_function:NN \langle seq \rangle \langle function \rangle
\end{verbatim}

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:

\begin{verbatim}
\sys_if_engine_xetex:TF ✯ \sys_if_engine_xetex:TF \langle true code \rangle \langle false code \rangle
\end{verbatim}

The underlining and italic of \(T\) indicates that three functions are available:

- \sys_if_engine_xetex:T
- \sys_if_engine_xetex:F
- \sys_if_engine_xetex:TF

Usually, the illustration will use the \(TF\) variant, and so both \langle true code \rangle and \langle false code \rangle will be shown. The two variant forms \(T\) and \(F\) take only \langle true code \rangle and \langle false code \rangle, respectively. Here, the star also shows that this function is expandable. With some minor exceptions, 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 ✯ \token_to_str:N \langle token \rangle
\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 \(\LaTeX\) or plain \(\TeX\). In these cases, the text will include an extra \texttt{\TeXhackers note} section:

\begin{verbatim}
\token_to_str:N ✯ \token_to_str:N \langle token \rangle
\end{verbatim}

The normal description text.

\TeXhackers note: Detail for the experienced \TeX or \LaTeX 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 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 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.

1.3 Formal language conventions which apply generally

As this is a formal reference guide for \LaTeX3 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 TF argument specification, the test if evaluated to give a logically TRUE or FALSE result. Depending on this result, either the ⟨true code⟩ or the ⟨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.

1.4 \TeX concepts not supported by \LaTeX3

The \TeX concept of an “\textbackslash outer” macro is not supported at all by \LaTeX3. As such, the functions provided here may break when used on top of \LaTeX2ε if \textbackslash outer tokens are used in the arguments.
Part II

Bootstrapping
Chapter 2

The \texttt{l3bootstrap} package

Bootstrap code

2.1 Using the \LaTeX3 modules

The modules documented in source3 are designed to be used on top of \LaTeX2ε and are loaded all as one with the usual \texttt{\usepackage{expl3}} or \texttt{\RequirePackage{expl3}} instructions.

As the modules use a coding syntax different from standard \LaTeX2ε it provides a few functions for setting it up.

\begin{verbatim}
\ExplSyntaxOn
⟨code⟩ \ExplSyntaxOff
\end{verbatim}

The \texttt{\ExplSyntaxOn} function switches to a category code regime in which spaces and new lines 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 \texttt{\ExplSyntaxOff} reverts to the document category code regime.

\textbf{\TeXXhackers note:} Spaces introduced by - behave much in the same way as normal space characters in the standard category code regime: they are ignored after a control word or at the start of a line, and multiple consecutive - are equivalent to a single one. However, - is not ignored at the end of a line.

\begin{verbatim}
\ProvidesExplPackage \ProvidesExplClass \ProvidesExplFile
\end{verbatim}

These functions act broadly in the same way as the corresponding \LaTeX2ε kernel functions \texttt{\ProvidesPackage}, \texttt{\ProvidesClass} and \texttt{\ProvidesFile}. However, they also implicitly switch \texttt{\ExplSyntaxOn} for the remainder of the code with the file. At the end of the file, \texttt{\ExplSyntaxOff} will be called to reverse this. (This is the same concept as \LaTeX2ε provides in turning on \texttt{\makeatletter} within package and class code.) The \texttt{⟨date⟩} should be given in the format \texttt{(year)/(month)/(day)} or in the ISO date format \texttt{(year)-(month)-(day)}. If the \texttt{⟨version⟩} is given then it will be prefixed with \texttt{v} in the package identifier line.
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 \TeX{} category codes and the \TeX{} 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}
Chapter 3

The \texttt{l3names} package
Namespace for primitives

3.1 Setting up the \LaTeX{}3 programming language

This module is at the core of the \LaTeX{}3 programming language. It performs the following tasks:

- defines new names for all \TeX{} primitives;
- emulate required primitives not provided by default in \texttt{Lua\TeX};
- switches to the category code régime for programming;

This module is entirely dedicated to primitives (and emulations of these), which should not be used directly within \LaTeX{}3 code (outside of “kernel-level” code). As such, the primitives are not documented here: \textit{The \TeX{}book}, \textit{\TeX{} by Topic} and the manuals for \texttt{pdf\TeX}, \texttt{Xe\TeX}, \texttt{Lua\TeX}, \texttt{\p\TeX} and \texttt{up\TeX} should be consulted for details of the primitives. These are named \texttt{\textbackslash tex\_\{name\}}; typically based on the primitive’s \texttt{\{name\}} in \texttt{pdf\TeX} and omitting a leading \texttt{pdf} when the primitive is not related to \texttt{pdf} output.
Part III
Programming Flow
Chapter 4

The \texttt{l3basics} package

Basic definitions

As the name suggests, 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.

4.1 No operation functions

\begin{itemize}
  \item \texttt{\prg_do_nothing}:
    An expandable function which does nothing at all: leaves nothing in the input stream after a single expansion.
  \item \texttt{\scan_stop}:
    A non-expandable function which does nothing. Does not vanish on expansion but produces no typeset output.
\end{itemize}

4.2 Grouping material

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.
\texttt{\textbackslash group\_insert\_after:N} \texttt{\textbackslash group\_insert\_after:N (token)}

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{\textbackslash 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{\textbackslash group\_end:} function or by a token with category code 2 (close-group), namely a \texttt{)}} if standard category codes apply.

\texttt{\textbackslash group\_show\_list:} \texttt{\textbackslash group\_log\_list:}

Display (to the terminal or log file) a list of the groups that are currently opened. This is intended for tracking down problems.

\textbf{T\textit{eX}hackers note:} This is a wrapper around the \texttt{\textbackslash showgroups} primitive.

4.3 Control sequences and functions

As \texttt{T\textit{eX}} 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 \texttt{(#1, #2, etc.)} is replaced the appropriate arguments absorbed by the function. In the following, \texttt{(code)} is therefore used as a shorthand for “replacement text”.

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

4.3.1 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 \texttt{\textbackslash 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 \texttt{(#1, #2, ...)}.

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

\texttt{set} Create a new function with the \texttt{set} scope, such as \texttt{\cs\_set:Npn}. The definition is restricted to the current \texttt{T\textit{eX}} group and does not result in an error if the function is already defined.

\texttt{gset} Create a new function with the \texttt{gset} scope, such as \texttt{\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.

\texttt{nopar} Create a new function with the \texttt{nopar} restriction, such as \texttt{\cs\_set\_nopar:Npn}. The parameter may not contain \texttt{\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 4.3.2 and 4.3.3 are primarily meant to define base functions only. Base functions can only have the following argument specifiers:

N and n No manipulation.

T and F Functionally equivalent to n (you are actually encouraged to use the family of `\prg_new_conditional:` functions described in Section 9.1).

p and 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 5.2.

4.3.2 Defining new functions using parameter text

\begin{verbatim}
\cs_new:Npn \cs_new:cpn \cs_new:Npx \cs_new:cpx
\cs_new_protected:Npn \cs_new_protected:cpn \cs_new_protected:Npx \cs_new_protected:cpx
\end{verbatim}

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>\cs_new:Npn</code></td>
<td>Creates \textit{(function)} to expand to \textit{(code)} as replacement text. Within the \textit{(code)}, the \textit{(parameters) (#1, #2, etc.)} will be replaced by those absorbed by the function. The definition is global and an error results if the \textit{(function)} is already defined.</td>
</tr>
<tr>
<td><code>\cs_new:cpn</code></td>
<td>Creates \textit{(function)} to expand to \textit{(code)} as replacement text. Within the \textit{(code)}, the \textit{(parameters) (#1, #2, etc.)} will be replaced by those absorbed by the function. When the \textit{(function)} is used the \textit{(parameters) absorbed cannot contain \par tokens. The definition is global and an error results if the \textit{(function)} is already defined.</td>
</tr>
<tr>
<td><code>\cs_new:Npx</code></td>
<td>Creates \textit{(function)} to expand to \textit{(code)} as replacement text. Within the \textit{(code)}, the \textit{(parameters) (#1, #2, etc.)} will be replaced by those absorbed by the function. The \textit{(function)} will not expand within an \textit{x}-type argument. The definition is global and an error results if the \textit{(function)} is already defined.</td>
</tr>
<tr>
<td><code>\cs_new:cpx</code></td>
<td>Creates \textit{(function)} to expand to \textit{(code)} as replacement text. Within the \textit{(code)}, the \textit{(parameters) (#1, #2, etc.)} will be replaced by those absorbed by the function. When the \textit{(function)} is used the \textit{(parameters) absorbed cannot contain \par tokens. The \textit{(function)} will not expand within an \textit{x}-type or \textit{e}-type argument. The definition is global and an error results if the \textit{(function)} is already defined.</td>
</tr>
</tbody>
</table>

15
\cs_set:Npn \cs_set:cpn \cs_set:Npx \cs_set:cpx
\cs_set_protected:Npn \cs_set_protected:cpn \cs_set_protected:Npx \cs_set_protected:cpx
\cs_gset:Npn \cs_gset:cpn \cs_gset:Npx \cs_gset:cpx

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

\cs_set_nopar:Npn \cs_set_nopar:cpn \cs_set_nopar:Npx \cs_set_nopar:cpx
\cs_set_protected_nopar:Npn \cs_set_protected_nopar:cpn \cs_set_protected_nopar:Npx \cs_set_protected_nopar:cpx
\cs_gset_nopar:Npn \cs_gset_nopar:cpn \cs_gset_nopar:Npx \cs_gset_nopar:cpx

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

\texttt{x}\texttt{-type or }	exttt{e}\texttt{-type argument.}

\cs_gset:Npn \cs_gset:cpn \cs_gset:Npx \cs_gset:cpx
\cs_gset_protected:Npn \cs_gset_protected:cpn \cs_gset_protected:Npx \cs_gset_protected:cpx
\cs_gset_nopar:Npn \cs_gset_nopar:cpn \cs_gset_nopar:Npx \cs_gset_nopar:cpx
\cs_gset_protected_nopar:Npn \cs_gset_protected_nopar:cpn \cs_gset_protected_nopar:Npx \cs_gset_protected_nopar:cpx
\cs_gset:Nopar \cs_gset:Nopar:cpn \cs_gset:Nopar:Npx \cs_gset:Nopar:Nopar:cpn
\cs_gset_protected:Nopar \cs_gset_protected:cpn \cs_gset_protected:Nopar:Npx \cs_gset_protected:Nopar:Nopar:cpn
\cs_gset_nopar:Nopar \cs_gset_nopar:cpn \cs_gset_nopar:Nopar:Npx \cs_gset_nopar:Nopar:Nopar:cpn

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

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

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

\cs_new:Nn \cs_new:nopar:Nn \cs_new_protected:Nn \cs_gset_protected_nopar:Npn

Globally sets \textit{function} to expand to \textit{code} as replacement text. Within the \textit{code}, the \textit{parameters} (\#1, \#2, 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 \textit{not} restricted to the current \TeX{} group level: the assignment is global. The \textit{function} will not expand within an x-type argument.

\cs_new:Nn \cs_new:nopar:Nn \cs_new_protected:Nn \cs_gset_protected_nopar:Npn

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, etc.) will be replaced by those absorbed by the function. The definition is global and an error results if the \textit{function} is already defined.

\cs_new:Nn \cs_new:nopar:Nn \cs_new_protected:Nn \cs_gset_protected_nopar:Npn

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, 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_new:Nn \cs_new:nopar:Nn \cs_new_protected:Nn \cs_gset_protected_nopar:Npn

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, 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:nopar:Nn \cs_new_protected_nopar:Nn \cs_gset_protected_nopar:Npn

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, 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_protected:Nn  \cs_set_protected:(cn|Nx|cx)
\cs_set_protected:Nn \langle function \rangle \{ \langle code \rangle \}

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the number of \langle parameters \rangle is detected automatically from the function signature. These \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 \texttt{\par} tokens. The assignment of a meaning to the \langle function \rangle is restricted to the current \TeX{} group level.

\cs_set_protected:Nn  \cs_set_protected:(cn|Nx|cx)
\cs_set_protected:Nn \langle function \rangle \{ \langle code \rangle \}

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

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the number of \langle parameters \rangle is detected automatically from the function signature. These \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 \texttt{\par} tokens. The \langle function \rangle will not expand within an \texttt{x}-type or \texttt{e}-type argument. The assignment of a meaning to the \langle function \rangle is restricted to the current \TeX{} group level.

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

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the number of \langle parameters \rangle is detected automatically from the function signature. These \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 global.

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

Sets \langle function \rangle to expand to \langle code \rangle as replacement text. Within the \langle code \rangle, the number of \langle parameters \rangle is detected automatically from the function signature. These \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 \texttt{\par} tokens. The assignment of a meaning to the \langle function \rangle is global.

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

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

uses the \textit{creator} function (which should have signature \texttt{Npn}, for example \texttt{cs_new:Npn}) to define a \textit{function} which takes \textit{number} arguments and has \textit{code} as replacement text. the \textit{number} of arguments is an integer expression, evaluated as detailed for \texttt{int_eval:n}.

4.3.4 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 \textit{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”.

Globally creates \textit{control sequence} and sets it to have the same meaning as \textit{control sequence} or \textit{token}. the second control sequence may subsequently be altered without affecting the copy.

Sets \textit{control sequence} to have the same meaning as \textit{control sequence} (or \textit{token}). the second control sequence may subsequently be altered without affecting the copy. the assignment of a meaning to the \textit{control sequence} is restricted to the current \TeX group level.

Globally sets \textit{control sequence} to have the same meaning as \textit{control sequence} (or \textit{token}). the second control sequence may subsequently be altered without affecting the copy. the assignment of a meaning to the \textit{control sequence} is \textit{not} restricted to the current \TeX group level: the assignment is global.
4.3.5 Deleting control sequences

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

\cs_undefine:N \langle control sequence \rangle

Sets \langle control sequence \rangle to be globally undefined.

4.3.6 Showing control sequences

\cs_meaning:N \langle control sequence \rangle

This function expands to the meaning of the \langle control sequence \rangle control sequence. For a macro, this includes the \langle replacement text \rangle.

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

\cs_show:N \langle control sequence \rangle

Displays the definition of the \langle control sequence \rangle on the terminal.

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

\cs_log:N \langle control sequence \rangle

Writes the definition of the \langle control sequence \rangle in the log file. See also \cs_show:N which displays the result in the terminal.

4.3.7 Converting to and from control sequences

\use:c \langle control sequence name \rangle

Expands the \langle control sequence name \rangle 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 \langle control sequence name \rangle must, when fully expanded, consist of character tokens, typically a mixture of category code 10 (space), 11 (letter) and 12 (other).

**\TeXhackers note:** Protected macros that appear in a \texttt{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.

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

and

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

would be equivalent to

\abc

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

\cs:w \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.

\TeXhacks note: These are the \TeX primitives \texttt{\csname} and \texttt{\endcsname}.

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

\cs:w a b c \cs_end:

and

\tl_new:N \l_my_tl
\tl_set:Nn \l_my_tl \{ a b c \}
\cs:w \tl_use:N \l_my_tl \cs_end:

would be equivalent to

\abc

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

\cs_to_str:N

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 not include the current escape token, contrarily to \texttt{\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.
4.4 Analysing control sequences

\cs_split_function:N \langle function \rangle

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

The next three functions decompose \TeX{} macros into their constituent parts: if the \langle token \rangle passed is not a macro then no decomposition can occur. In the latter case, all three functions leave \texttt{\\bec\bec}\ in the input stream.

\cs_prefix_spec:N \langle token \rangle

If the \langle 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

\texttt{\cs_set:Npn \next:nn \#1\#2 \{ x \#1-y \#2 \}}
\texttt{\cs_prefix_spec:N \next:nn}

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

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

\cs_argument_spec:N \langle token \rangle

If the \langle 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

\texttt{\cs_set:Npn \next:nn \#1\#2 \{ x \#1 \#2 \}}
\texttt{\cs_argument_spec:N \next:nn}

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

\TeX{}hackers note: If the argument specification contains the string -\rightarrow, then the function produces incorrect results.
\cs_replacement_spec:N \langle \text{token} \rangle

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-y \#2 \}
\cs_replacement_spec:N \next:nn
\end{verbatim}

leaves \texttt{x#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.

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

\section{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).

\begin{verbatim}
\use:n \langle \text{group}_1 \rangle
\use:nn \langle \text{group}_1 \rangle \{ \langle \text{group}_2 \rangle \}
\use:nnn \langle \text{group}_1 \rangle \{ \langle \text{group}_2 \rangle \} \{ \langle \text{group}_1 \rangle \}
\use:nnn \langle \text{group}_1 \rangle \{ \langle \text{group}_2 \rangle \} \{ \langle \text{group}_1 \rangle \} \{ \langle \text{group}_4 \rangle \}
\end{verbatim}

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

\begin{verbatim}
\use:nn { abc } \{ { def } }\n\end{verbatim}

results in the input stream containing

\begin{verbatim}
abc { def }\n\end{verbatim}

\textit{i.e.} only the outer braces are removed.

\textbf{\LaTeX}hackers note: The \texttt{\use:n} function is equivalent to \LaTeX{}2e’s \texttt{\@firstofone}. 
\use_i:nn \use_i:nn \{\langle \text{arg}_1 \rangle \}\{\langle \text{arg}_2 \rangle \}\{\langle \text{arg}_3 \rangle \}

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.

\textbf{\LaTeX{}hackers note:} These are equivalent to \LaTeX{}2ε’s \texttt{\@firstoftwo} and \texttt{\@seconddoftwo}.

\use_iii:nn \use_iii:nn \{\langle \text{arg}_1 \rangle \}\{\langle \text{arg}_2 \rangle \}\{\langle \text{arg}_3 \rangle \}\{\langle \text{arg}_4 \rangle \}

These functions absorb three arguments from the input stream. The function \use_iii:nn discards the second and third arguments, and leaves the content of the first argument in the input stream. \use_ii:nn and \use_ii: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.

\use_i:nnnn \use_i:nnnn \{\langle \text{arg}_1 \rangle \}\{\langle \text{arg}_2 \rangle \}\{\langle \text{arg}_3 \rangle \}\{\langle \text{arg}_4 \rangle \}\{\langle \text{arg}_5 \rangle \}

These functions absorb four arguments from the input stream. The function \use_i:nnn 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.

\use_ii:nnn \use_ii:nnn \{\langle \text{arg}_1 \rangle \}\{\langle \text{arg}_2 \rangle \}\{\langle \text{arg}_3 \rangle \}\{\langle \text{arg}_4 \rangle \}

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:

\begin{verbatim}
\use_i_ii:nnn { abc } { { def } } { ghi }
\end{verbatim}

results in the input stream containing

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

\textit{i.e.} the outer braces are removed and the third group is removed.

\use_ii_i:nn \use_ii_i:nn \{\langle \text{arg}_1 \rangle \}\{\langle \text{arg}_2 \rangle \}

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.
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 \texttt{n} arguments may be an unbraced single token (i.e. an \texttt{N} argument).

\textbf{\LaTeXhackers note:} These are equivalent to \LaTeX\texttt{2e}'s \texttt{@gobble}, \texttt{@gobbletwo}, etc.

\begin{verbatim}
\use:⋆\x
\end{verbatim}

\texttt{\use:⋆\x} is a wrapper around the primitive \texttt{\expanded} where it is available: it requires two expansions to complete its action. When \texttt{\expanded} is not available this function is very slow.

\section*{4.5.1 Selecting tokens from delimited arguments}

A different kind of function for selecting tokens from the token stream are those that use delimited arguments.

\begin{verbatim}
\use:⋆\x
\end{verbatim}

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

\textbf{\LaTeXhackers note:} \texttt{\use:⋆\x} is a wrapper around the primitive \texttt{\expanded} where it is available: it requires two expansions to complete its action. When \texttt{\expanded} is not available this function is very slow.

\section*{4.6 Predicates and conditionals}

\LaTeX\texttt{3} 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 (true code) or the (false code). These arguments are denoted with T and F, respectively. An example would be

\cs_if_free:cTF {abc} {⟨true code⟩} {⟨false code⟩}

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 ⟨true code⟩ and/or ⟨false code⟩ 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,

\cs_if_free_p:N

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

\bool_if:nTF {
    \cs_if_free_p:N \l_tmpz_tl || \cs_if_free_p:N \g_tmpz_tl
} {⟨true code⟩} {⟨false code⟩}

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.

\c_true_bool \c_false_bool    Constants that represent true and false, respectively. Used to implement predicates.
4.6.1 Tests on control sequences

\cs_if_eq_p:NN
\cs_if_eq:NNTF

Compares the definition of two \emph{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}.

\cs_if_exist_p:N
\cs_if_exist:NTF
\cs_if_free_p:N
\cs_if_free:NTF

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

\cs_if_meaning:w
\reverse_if:N
\if_meaning:w
\reverse_if:N

\TeXhacksnote{These are equivalent to their corresponding \TeX\ primitive conditionals; \texttt{\reverse_if:N} is \TeXX's \texttt{\unless}.}

4.6.2 Primitive conditionals

The \TeXX 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_false:
\else:
\fi:
\reverse_if:N

\TeXhacksnote{These \TeXX hackers note: This is \TeX's \texttt{\ifx}.}
\if:w *
\if_charcode:w *
\if_catcode:w *
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:w *
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_vertical: *
\if_mode_math: *
\if_mode_inner: *
4.7 Starting a paragraph

\mode_leave_vertical: 
\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 \TeX{}2\e
\leavevmode approach, no box is used by the method implemented here.
4.8 Debugging support

\debug_on:n
\debug_off:n

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 \langle list \rangle are

- **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;
- **check-expressions** that checks integer, dimension, skip, and muskip expressions are not terminated prematurely;
- **deprecation** that makes soon-to-be-deprecated commands produce errors;
- **log-functions** that logs function definitions;
- **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\ package mode loaded with \enabledebug or another option implying it.

\debug_suspend: \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 l3coffins.
Chapter 5

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 \texttt{exp}.

Not all possible variations are implemented for every base function. Instead only those that are used within the \LaTeX{} 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.

5.1 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 \texttt{\exp_\ldots}. They all look alike, an example would be \texttt{\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 \texttt{\exp_args:NNo} expands the content of third argument once before any expansion of the first and second arguments. If \texttt{\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 \texttt{\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 \texttt{\exp\_} function followed by the desired base function, \textit{e.g.}

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

results in the definition of \texttt{\seq_gpush:No}
Providing variants in this way in style files is safe as the \texttt{\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 \texttt{\cs_generate_variant:Nn}, described next.

### 5.2 Methods for defining variants

We recall the set of available argument specifiers.

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

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

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

This means the \textit{parent} of a \textit{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 \texttt{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 a \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.

### 5.3 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 \textbackslash 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 token in the argument it receives: it is conceptually similar to

\begin{verbatim}
\exp_after:wN \texttt{<base function>} \exp_after:wN \{ \texttt{<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:N \tl_tmpa_tl \{ \{ \texttt{\g_tmpb_tl} \} \}
\end{verbatim}
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.

5.4 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:cc ⋆
\exp_args:Nc ⟨function⟩ {⟨tokens⟩}
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⟩.

\textbf{\TeXhackers note:} Protected macros that appear in a 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.
This function absorbs two arguments (the ⟨function⟩ name and the ⟨tokens⟩). The ⟨tokens⟩ are expanded once, and 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.

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

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. This control sequence should be the name of a ⟨variable⟩. The content of the ⟨variable⟩ are recovered and placed inside braces into the input stream after reinsertion of the ⟨function⟩. Thus the ⟨function⟩ may take more than one argument: all others are left unchanged.

\TeXhackers 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.

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.

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

### 5.5 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 are not expandable due to their \textit{x}-type argument.
5.6 Manipulating three arguments

\exp_args:NNNo \langle \langle \langle \text{tokens} \rangle \rangle \rangle

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, etc.

\exp_args:NNcf \langle \langle \langle \langle \text{tokens} \rangle \rangle \rangle \rangle

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, etc. These functions need slower processing.

\exp_args:NNnx \langle \langle \langle \langle \text{tokens} \rangle \rangle \rangle \rangle

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, etc.
5.7 Unbraced expansion

\exp_last_unbraced:Nx \{token\} \{\tokens_1\} \{\tokens_2\}

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,
\exp_last_unbraced:Nf \foo_bar:w \{ \} \q_stop leads to an infinite loop, as the quark is f-
expanded.

\exp_last_unbraced:Nx
\exp_last_two_unbraced:Noo
\exp_after:wN

\exp_last_unbraced:Nx \{function\} \{\tokens\}

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

\exp_last_two_unbraced:Noo \{token\} \{\tokens_1\} \{\tokens_2\}

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 \{token\} \{token_2\}

Carries out a single expansion of \{token_2\} (which may consume arguments) prior to the
expansion of \{token_1\}. If \{token_2\} has no expansion (for example, if it is a character) then
it is left unchanged. It is important to notice that \{token_1\} may be any single token, in-
cluding 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 \exp_arg:N function.

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

5.8 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 ‘expand-
able’ since they themselves disappear after the expansion has completed.

\texttt{\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 (\texttt{\cs_new:Npx}), a macro parameter introduces an argument even if it appears as \exp_not:N \# 1. This differs from \exp_not:n.

\texttt{\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.

\texttt{\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 \texttt{\exp_not:n} \texttt{\exp_not:n} primitive. In an x-expanding definition (\texttt{\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 \#.

\texttt{\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.

\texttt{\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.
\texttt{\textbackslash exp\_not:v} \texttt{(\{tokens\})} 

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

\textbf{\LaTeX{}hackers note:} Protected macros that appear in a 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.

\texttt{\textbackslash exp\_not:e} \texttt{(\{tokens\})} 

Expands \texttt{\{tokens\}} 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 \texttt{\exp_not:n}. This is very rarely useful but is provided for consistency.

\texttt{\textbackslash exp\_not:f} \texttt{(\{tokens\})} 

Expands \texttt{\{tokens\}} 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 \texttt{\exp_not:n}.

\texttt{\exp_stop_f:} \texttt{(\{tokens\} \exp_stop_f: \{more tokens\})} 

This function terminates an f-type expansion. Thus if a function \texttt{\foo_bar:f} starts an f-type expansion and all of \texttt{\{tokens\}} are expandable \texttt{\exp_stop_f:} terminates the expansion of tokens even if \texttt{\{more tokens\}} 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 (\texttt{\textbackslash \space}).

5.9 Controlled expansion

The expl3 language makes all efforts to hide the complexity of \TeX{} expansion from the programmer by providing concepts that evaluate/expand arguments of functions prior to calling the “base” functions. Thus, instead of using many \texttt{\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 \TeX{} 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 \texttt{\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 \texttt{\{expandable\-tokens\}} as that will break badly if unexpandable tokens are encountered in that place!
\exp:w \exp_end:
\exp:w \langle expandable-tokens \rangle \exp_end:
Expands \langle expandable-tokens \rangle until reaching \exp_end: at which point expansion stops. The full expansion of \langle expandable tokens \rangle has to be empty. If any token in \langle 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 expandable-tokens \rangle rather than being on the same expansion level than \exp:w, e.g., you may see code such as

\exp:w \@@_case:NnTF #1 {#2} { } { }
where somewhere during the expansion of \@@_case:NnTF the \exp_end: gets generated.

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

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

The full expansion of \langle expandable-tokens \rangle has to be empty. If any token in \langle 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 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 #2 }
where the \exp_after:wN triggers an $f$-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 $f$-expansion).

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

\exp:w \langle expandable-tokens \rangle \exp_end:
can be alternatively achieved through an $f$-type expansion by using \exp_stop_f:, i.e.

\exp:w \exp_end_continue_f:w \langle 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.

4Due to the implementation you might get the character in position 0 in the current font (typically ‘’ in the output without any error message!)

5In this particular case you may get a character into the output as well as an error message.
The difference to \texttt{\exp_end_continue_f:w} is that we first we pick up an argument which is then returned to the input stream. If \texttt{(further-tokens)} 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 \texttt{f-type} expansion.

\section{5.10 Internal functions}

\begin{verbatim}
\cs_new:Npn \exp: \exp_end_continue_f:w \exp_end_continue_f:nw \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w \exp_end_continue_f:w

\cs_new:Npn \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno \exp_last_unbraced:Nno

\end{verbatim}

Internal forms for the base expansion types. These names do \textit{not} conform to the general \LaTeX\3 approach as this makes them more readily visible in the log and so forth. They should not be used outside this module.

Internal forms for the expansion types which leave the terminal argument unbraced. These names do \textit{not} conform to the general \LaTeX\3 approach as this makes them more readily visible in the log and so forth. They should not be used outside this module.
Chapter 6

The \texttt{l3sort} package

Sorting functions

6.1 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 \{-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 \texttt{⟨comparison code⟩} 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 \texttt{⟨comparison code⟩} consisting only of \texttt{\sort_return_swapped:} reverses the list (in a fairly inefficient way).

\textbf{\LaTeX{}hackers note:} The current implementation is limited to sorting approximately 20000 items (40000 in \LaTeX{}), depending on what other packages are loaded.

Internally, the code from \texttt{l3sort} stores items in \texttt{\toks} registers allocated locally. Thus, the \texttt{⟨comparison code⟩} 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.
\texttt{\textbackslash sort\_return\_same}: \texttt{\textbackslash seq\_sort:}\texttt{Nn} (seq var)
\texttt{\textbackslash sort\_return\_swapped}: { ... \texttt{\textbackslash sort\_return\_same}: or \texttt{\textbackslash sort\_return\_swapped}: ... }

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{\textbackslash sort\_return\_...} 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.
Chapter 7

The \texttt{l3tl-analysis} package: Analysing token lists

This module provides functions that are particularly useful in the \texttt{l3regex} module for mapping through a token list one \texttt{token} at a time (including begin-group/end-group tokens). For \texttt{tl_analysis_map_inline:Nn} or \texttt{tl_analysis_map_inline:nn}, the token list is given as an argument; the analogous function \texttt{peek_analysis_map_inline:n} documented in \texttt{l3token} finds tokens in the input stream instead. In both cases the user provides \texttt{inline code} that receives three arguments for each \texttt{token}:

- \texttt{\langle tokens \rangle}, which both o-expand and x-expand to the \texttt{token}. The detailed form of \texttt{\langle tokens \rangle} may change in later releases.
- \texttt{\langle char code \rangle}, a decimal representation of the character code of the \texttt{\langle token \rangle}, $-1$ if it is a control sequence.
- \texttt{\langle catcode \rangle}, a capital hexadecimal digit which denotes the category code of the \texttt{\langle token \rangle} (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). This can be converted to an integer by writing \texttt{\langle \langle catcode \rangle \rangle}.

In addition, there is a debugging function \texttt{tl_analysis_show:n}, very similar to the \texttt{ShowTokens} macro from the \texttt{ted} package.

\begin{verbatim}
\tl_analysis_show:n \langle token list \rangle
\tl_analysis_log:n \langle token list \rangle
\end{verbatim}

Displays to the terminal (or log) the detailed decomposition of the \texttt{\langle token list \rangle} 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 \langle token list \rangle \langle inline function \rangle
\tl_analysis_map_inline:Nn \langle token list \rangle \langle inline function \rangle
\end{verbatim}

Applies the \texttt{\langle inline function \rangle} to each individual \texttt{\langle token \rangle} in the \texttt{\langle token list \rangle}. The \texttt{\langle inline function \rangle} receives three arguments as explained above. As all other mappings the mapping is done at the current group level, \textit{i.e.} any local assignments made by the \texttt{\langle inline function \rangle} remain in effect after the loop.
Chapter 8

The \texttt{l3regex} package: Regular expressions in \TeX

The \texttt{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 \texttt{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 \texttt{\l_my_tl} holds the text “This cat.”, where the first occurrence of “at” was replaced by “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 \texttt{\w} sequence represents any “word” character, and + indicates that the \texttt{\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, \texttt{\0} denotes the full match (here, a word). The command \texttt{\emph} is inserted using \texttt{\c{emph}}, and its argument \texttt{\0} is put between braces \texttt{\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 \texttt{\regex_const:Nn}. For example,

\begin{verbatim}
\regex_const:Nn \c_foo_regex { \c{begin} \cB. \c[^BE].* \cE. }
\end{verbatim}

stores in \texttt{\c_foo_regex} a regular expression which matches the starting marker for an environment: \texttt{\begin}, followed by a begin-group token (\texttt{\cB.}), then any number of tokens which are neither begin-group nor end-group character tokens (\texttt{\c[^BE].*}), ending with an end-group token (\texttt{\cE.}). As explained in the next section, the parentheses “capture” the result of \texttt{\c[^BE].*}, giving us access to the name of the environment when doing replacements.
8.1 Syntax of regular expressions

8.1.1 Regex examples

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 \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+\d+** matches an explicit integer with at most one sign.

- **[\+\-\_]\*\d+\d+\d+\d+** matches an explicit integer with any number of + and – signs, with spaces allowed except within the mantissa, and surrounded by spaces.

- **[\+\-\_]\*(\d+\d+\d+\d+)\d+\d+** matches an explicit integer or decimal number; using [.,] instead of \d+ would allow the comma as a decimal marker.

- **[\+\-\_]\*(\d+\d+\d+\d+)\d+\d+\d+\d+\d+** matches an explicit dimension with any unit that \TeX\ knows, where (?i) means to treat lowercase and uppercase letters identically.

- **[\+\-\_]\*((?i)nan|inf|(?i)pt|in|\{cem\}m|ex|\{bs\}p|dn|\{pcn\}c\d+\d+** matches an explicit floating point number or the special values nan and inf (with signs and spaces allowed).

- **[\+\-\_]\*((?i)nan|inf|(?i)pt|in|\{cem\}m|ex|\{bs\}p|dn|\{pcn\}c\d+\d+** 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 \K) everything between the end of the previous match (\G) and what is matched by the rest of the regular expression; this is useful in \regex_replace_all:nnN when the goal is to extract matches or submatches in a finer way than with \regex_extract_all:nnN.

While it is impossible for a regular expression to match only integer expressions, [\+\-\j\d+\d+\d+\d+] among other things all valid integer expressions (made only with explicit integers). One should follow it with further testing.
8.1.2 Characters in regular expressions

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 \d 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 \(, \), \?, \.);
- 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).

8.1.3 Characters classes

Character types.

. A single period matches any token.
\d Any decimal digit.
\h Any horizontal space character, equivalent to [\ \^-][\^-]-[\^-][\^-]-[\^-][\^-][\^-][\^-][\^-][\^-][\^-][\^-][\^-].
\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.

\\ 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 \langle name\rangle, which can be alnum, alpha, ascii, blank, cntrl, digit, graph, lower, print, punct, space, upper, word, or xdigit.

[::^\langle name\rangle:] 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).

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.

8.1.4 Structure: alternatives, groups, repetitions

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.

For greedy quantifiers the regex code will first investigate matches that involve as many repetitions as possible, while for lazy quantifiers it investigates matches with as few repetitions as possible first.

Alternation and capturing groups.
A|B|C Either one of A, B, or C, investigating A first.
(... ) Capturing group.
(??:...) Non-capturing group.

Either one of A, B, or C, investigating A first.

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 \{ai\} and \{aa\}, but they are trimmed by the use of \K. The \K command does not affect capturing groups: for instance,

\regex_extract_once:nnN { (. \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.

8.1.5 Matching exact tokens

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;
• 0 for others; and
• A for active characters.

The \c escape sequence is used as follows.

\c{⟨regex⟩} A control sequence whose csname matches the ⟨regex⟩, anchored at the beginning and end, so that \c{begin} matches exactly \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 X (any of CBEMTPUDSLOA. For instance, \cL[A-Z\d] matches uppercase letters and digits of category code letter, \cC. matches any control sequence, and \cO(abc) matches abc where each character has category other.

\c[XYZ] Applies to the next object, and forces it to only match tokens with category X, Y, or Z (each being any of 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 X, Y, or Z (each being any of 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, [\cO\d \c[LO][A-F]] matches what \pX 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, \cL(ab\cO\*cd) matches ab*cd 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{⟨var name⟩} matches the exact contents (both character codes and category codes) of the variable \⟨var name⟩, which are obtained by applying \exp_not:v {⟨var name⟩} at the time the regular expression is compiled. Within a \c{...} control sequence matching, the \u escape sequence only expands its argument once, in effect performing \tl_to_str:v. Quantifiers are supported.

The \ur escape sequence allows to insert the contents of a regex variable into a larger regular expression. For instance, A\ur{1_tmba_regex}D matches the tokens A and D separated by something that matches the regular expression \l_tmba_regex. This behaves as if a non-capturing group were surrounding \l_tmba_regex, and any group
contained in \_tmpa_regex is converted to a non-capturing group. Quantifiers are supported.

For instance, if \_tmpa_regex has value B|C, then A\ur{l\_tmpa_regex}D is equivalent to A(?:B|C)D (matching ABD or ACD) and not to AB|CD (matching AB or CD). To get the latter effect, it is simplest to use \TeX’s expansion machinery directly: if \_mymodule_BC_tl contains B|C then the following two lines show the same result:

\begin{verbatim}
\regex_show:n { A \u{l\_mymodule_BC_tl} D }
\regex_show:n { A B | C D }
\end{verbatim}

8.1.6 Miscellaneous

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.

\$ 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.

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.

8.2 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, \3, ... \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,

\begin{verbatim}
\tl_set:Nn \l_my_tl { Hello,-world! }
\regex_replace_all:nnN { ([er]?l|o) . } { (\0--\1) } \l_my_tl
\end{verbatim}

results in \l_my_tl holding H(ell--el)(o--o) w(or--o)(ld--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 \n-th submatch is empty if there are fewer than \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.

By default, the category code of characters inserted by the replacement are determined by the prevailing category code regime at the time where the replacement is made, with two exceptions:

• space characters (with character code 32) inserted with \c\ or \x20 or \x{20} have category code 10 regardless of the prevailing category code regime;

• if the category code would be 0 (escape), 5 (newline), 9 (ignore), 14 (comment) or 15 (invalid), it is replaced by 12 (other) instead.

The escape sequence \c allows to insert characters with arbitrary category codes, as well as control sequences.

\cX(\ldots) Produces the characters “…” with category \texttt{X}, which must be one of \texttt{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\{\langle text\rangle\} Produces the control sequence with csname \langle text\rangle. The \langle text\rangle may contain references to the submatches \0, \1, and so on, as in the example for \u below.

The escape sequence \u\{\langle\var name\rangle\} allows to insert the contents of the variable with name \langle\var name\rangle directly into the replacement, giving an easier control of category codes. When nested in \c\{\ldots\} and \u\{\ldots\} 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,

\begin{verbatim}
\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
\end{verbatim}

results in \l_my_tl holding first,\emph{second},first,first.

Regex replacement is also a convenient way to produce token lists with arbitrary category codes. For instance
\tl_clear:N \l_tmpa_tl
\regex_replace_all:nnN \{ \} { \cU\% \cA\~ } \l_tmpa_tl

results in \l_tmpa_tl containing the percent character with category code 7 (superscript) and an active tilde character.

8.3 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{l3regex} module’s functions can be given their regular expression argument either as an explicit string or as a compiled regular expression.

\regex_new:N \l_my_regex
\regex_set:Nn \l_my_regex { my\ (simple\ )? \texttt{reg(ular\ expression)} }

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.

8.4 Matching

All regular expression functions are available in both :n and :N variants. The former require a “standard” regular expression, while the later require a compiled expression as generated by \texttt{\regex_set:Nn}.
\regex_match:nnTF \regex_match:Nn

Tests whether the \langle regular expression \rangle matches any part of the \langle token list \rangle. For instance,

\begin{verbatim}
\regex_match:nnTF { b [cde]* } { abecdcx } { TRUE } { FALSE }
\regex_match:nnTF { [b-dq-w] } { example } { TRUE } { FALSE }
\end{verbatim}

leaves TRUE then FALSE in the input stream.

\regex_count:nnN \regex_count:Nn

Sets \langle int var \rangle within the current \TeX{} group level equal to the number of times \langle regular expression \rangle appears in \langle token list \rangle. 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 \l_foo_int taking the value 5.

### 8.5 Submatch extraction

\regex_extract_once:nnN \regex_extract_once:nnN

Finds the first match of the \langle regular expression \rangle in the \langle token list \rangle. If it exists, the match is stored as the first item of the \langle seq var \rangle, and further items are the contents of capturing groups, in the order of their opening parenthesis. The \langle seq var \rangle is assigned locally. If there is no match, the \langle seq var \rangle is cleared. The testing versions insert the \langle true code \rangle into the input stream if a match was found, and the \langle false code \rangle otherwise.

For instance, assume that you type

\begin{verbatim}
\regex_extract_once:nnTF { \A(La)?TeX(!*)\Z } { LaTeX!!! } \l_foo_seq
{ true } { false }
\end{verbatim}

Then the regular expression (anchored at the start with \A and at the end with \Z) must match the whole token list. The first capturing group, (La)?, matches La, and the second capturing group, (!*), matches !!!!!. Thus, \l_foo_seq contains as a result the items \{LaTeX!!!\}, \{La\}, and \{!!!\}, and the true branch is left in the input stream. Note that the \textit{n}-th item of \l_foo_seq, as obtained using \texttt{\seq_item:Nn}, correspond to the submatch numbered \((n - 1)\) in functions such as \texttt{\regex_replace_once:nnN}.
\regex_extract_all:nnN \{ \regex \} \{ \{ \token \} \} \{ \seq \} \{ \{ \true \} \} \{ \{ \false \} \} \\
\regex_extract_all:nnNTF \{ \regex \} \{ \{ \token \} \} \{ \seq \} \{ \{ \true \} \} \{ \false \} \\
\regex_extract_all:NN \{ \regex \} \{ \{ \token \} \} \{ \seq \} \\
\regex_extract_all:NNTF \{ \regex \} \{ \{ \token \} \} \{ \seq \} \{ \{ \true \} \} \{ \false \} \\
\regex_extract_all:NN \{ \regex \} \{ \{ \token \} \} \{ \seq \} \\
\regex_extract_all:NNTF \{ \regex \} \{ \{ \token \} \} \{ \seq \} \{ \{ \true \} \} \{ \false \} \\
\regex_extract_all:nnN \{ \{ \w+ \} \} \{ \{ \texttt{Hello,\,-\,world!} \} \} \l\_\_foo_seq \\
\{ \true \} \{ \false \} \\
Finds all matches of the \langle regular expression \rangle in the \langle token list \rangle, and stores all the sub-match information in a single sequence (concatenating the results of multiple \regex_extract_once:nnN calls). The \langle seq var \rangle is assigned locally. If there is no match, the \langle seq var \rangle is cleared. The testing versions insert the \langle true code \rangle into the input stream if a match was found, and the \langle false code \rangle otherwise. For instance, assume that you type \\
\regex_extract_all:nnNTF \{ \w+ \} \{ \{ \texttt{Hello,\,-\,world!} \} \} \l\_\_foo_seq \\
\{ \true \} \{ \false \} \\
Then the regular expression matches twice, the resulting sequence contains the two items \{Hello\} and \{world\}, and the \true branch is left in the input stream. \\
\regex_split:nnN \{ \{ \regular expression \} \} \{ \{ \token \} \} \{ \seq \} \\
\regex_split:nnNTF \{ \{ \regular expression \} \} \{ \{ \token \} \} \{ \seq \} \\
\regex_split:NN \{ \{ \regular expression \} \} \{ \{ \token \} \} \{ \seq \} \\
\regex_split:NNTF \{ \{ \regular expression \} \} \{ \{ \token \} \} \{ \seq \} \\
\regex_split:nnN \{ \{ \/ \} \} \{ \{ \texttt{the\,\,path\,\,for\,\,this\,\,file.tex} \} \} \l\_\_path_seq \\
\{ \true \} \{ \false \} \\
Splits the \langle token list \rangle into a sequence of parts, delimited by matches of the \langle regular expression \rangle. If the \langle regular expression \rangle has capturing groups, then the token lists that they match are stored as items of the sequence as well. The assignment to \langle seq var \rangle is local. If no match is found the resulting \langle seq var \rangle has the \langle token list \rangle as its sole item. If the \langle regular expression \rangle matches the empty token list, then the \langle token list \rangle is split into single tokens. The testing versions insert the \langle true code \rangle into the input stream if a match was found, and the \langle false code \rangle otherwise. For example, after \\
\seq_new:N \l\_\_path_seq \\
\regex_split:nnNTF \{ \{ \/ \} \} \{ \{ \texttt{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. \\
8.6 Replacement \\
\regex_replace_once:nnN \{ \{ \regular expression \} \} \{ \{ \replacement \} \} \{ \tl var \} \\
\regex_replace_once:nnNTF \{ \{ \regular expression \} \} \{ \{ \replacement \} \} \{ \tl var \} \{ \{ \true code \} \} \{ \{ \false code \} \} \\
Searches for the \langle regular expression \rangle in the \langle token list \rangle and replaces the first match with the \langle replacement \rangle. The result is assigned locally to \langle tl var \rangle. In the \langle replacement \rangle, \\0 represents the full match, \\1 represent the contents of the first capturing group, \\2 of the second, etc. \\
56
\regex_replace_all:nnN \regex_replace_all:nnNTF \regex_replace_all:NN \regex_replace_all:NNTTF
\regex_replace_all:nnN \regex_replace_all:nnNTF \regex_replace_all:NN \regex_replace_all:NNTTF

Replaces all occurrences of the \textit{regular expression} in the \textit{token list} by the \textit{replacement},
where \texttt{\0} represents the full match, \texttt{\1} represent the contents of the first capturing group,
\texttt{\2} of the second, \textit{etc}. Every match is treated independently, and matches cannot overlap.
The result is assigned locally to \texttt{\tl var}.

8.7 Constants and variables

\l_tmpa_regex \l_tmpb_regex
\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 \LaTeX3-defined function. However, they may be overwritten by
other non-kernel code and so should only be used for short-term storage.

Scratch regex 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.

8.8 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\texttt{\s?}}.
  Code improvements to come.
- Shift arrays so that the useful information starts at position 1.
- Only build \texttt{\c{...}} once.
- Use arrays for the left and right state stacks when compiling a regex.
• Should \_regex_action_free_group:n only be used for greedy \{n,\} quantifier? (I think not.)

• Quantifiers for \u and assertions.

• When matching, keep track of an explicit stack of curr_state and curr_submatches.

• If possible, when a state is reused by the same thread, kill other subthreads.

• Use an array rather than \l__regex_balance_tl to build the function \__regex_replacement_balance_one_match:n.

• Reduce the number of epsilon-transitions in alternatives.

• Optimize simple strings: use less states (abcade should give two states, for abc and ade). [Does that really make sense?]  

• Optimize groups with no alternative.

• Optimize states with a single \_regex_action_free:n.

• Optimize the use of \__regex_action_success: by inserting it in state 2 directly instead of having an extra transition.

• Optimize the use of \int_step... 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.
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 \textasciitilde x.

- Comments: \TeX\ already has its own \textasciitilde x.

- \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: Xe\TeX\ and Lua\TeX\ serve us characters directly, and splitting those into bytes is tricky, encoding dependent, and most likely not useful anyways.
Chapter 9

The \texttt{l3prg} package

Control structures

Conditional processing in \LaTeXe 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 \texttt{true} or \texttt{false}. For example, the function \texttt{\cs_if_free_p:N} checks whether the control sequence given as its argument is free and then returns the boolean \texttt{true} or \texttt{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 \texttt{N}) and then executes either \texttt{true} or \texttt{false} depending on the result.

\textbf{L3\TeX hackers note}: The arguments are executed after exiting the underlying \texttt{\if...\fi:} structure.

9.1 Defining a set of conditional functions

These functions create a family of conditionals using the same \texttt{(code)} to perform the test created. Those conditionals are expandable if \texttt{(code)} is. The \texttt{new} versions check for existing definitions and perform assignments globally (cf. \texttt{\cs_new:Npn}) whereas the \texttt{set} versions do no check and perform assignments locally (cf. \texttt{\cs_set:Npn}). The conditionals created are dependent on the comma-separated list of \texttt{(conditions)}, which should be one or more of \texttt{p, T, F and TF}. 

\texttt{\prg_new_conditional:Npnn} \texttt{\prg_set_conditional:Npnn} \texttt{\prg_new_conditional:Nnn} \texttt{\prg_set_conditional:Nnn} 
\texttt{\prg_new_conditional:Nn} \texttt{\prg_set_conditional:Nn} 
\texttt{\prg_new_conditional:Nn} \texttt{\prg_set_conditional:Nn}
These functions create a family of protected conditionals using the same \{\texttt{code}\} to perform the test created. The \{\texttt{code}\} does not need to be expandable. The \texttt{new} version check for existing definitions and perform assignments globally (cf. \texttt{cs_new:Npn}) whereas the \texttt{set} version do not (cf. \texttt{cs_set:Nnn}). The conditionals created are depended on the comma-separated list of \{\texttt{conditions}\}, which should be one or more of \texttt{T}, \texttt{F} and \texttt{TF} (not \texttt{p}).

The conditionals are defined by \texttt{prg_new_conditional:Npnn} and friends as:

- \texttt{\{name\}_p:\{arg spec\}} — a predicate function which will supply either a logical \texttt{true} or logical \texttt{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.

- \texttt{\{name\}:(arg spec)T} — a function with one more argument than the original \{arg spec\} demands. The \{true branch\} code in this additional argument will be left on the input stream only if the test is \texttt{true}.

- \texttt{\{name\}:(arg spec)F} — a function with one more argument than the original \{arg spec\} demands. The \{false branch\} code in this additional argument will be left on the input stream only if the test is \texttt{false}.

- \texttt{\{name\}:(arg spec)TF} — a function with two more argument than the original \{arg spec\} demands. The \{true branch\} code in the first additional argument will be left on the input stream if the test is \texttt{true}, while the \{false branch\} code in the second argument will be left on the input stream if the test is \texttt{false}.

The \{\texttt{code}\} of the test may use \{parameters\} as specified by the second argument to \texttt{prg_set_conditional:Npnn}: this should match the \{argument specification\} 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 \{\texttt{code}\}, 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:

\begin{verbatim}
pagr_set_conditional:Npnn \foo_if_bar:NN #1#2 { p , T , TF } {
  \if_meaning:w \l_tmpa_tl #1
  \prg_return_true:
  \else:
  \if_meaning:w \l_tmpa_tl #2
  \prg_return_true:
  \else:
  \prg_return_false:
  \fi:
  \fi:
}
\end{verbatim}
This defines the function \foo_if_bar_p:NN, \foo_if_bar:NNTF and \foo_if_bar:NNT but not \foo_if_bar:NNF (because F is missing from the \{conditions\} list). The return statements take care of resolving the remaining \else: and \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.

\begin{verbatim}
\prg_new_eq_conditional:NNn \langle name1 \rangle:\langle arg spec1 \rangle \langle name2 \rangle:\langle arg spec2 \rangle {\langle conditions \rangle}
\end{verbatim}

These functions copy a family of conditionals. The new version checks for existing definitions (cf. \cs_new_eq:NN) whereas the set version does not (cf. \cs_set_eq:NN). The conditionals copied are depended on the comma-separated list of \langle conditions\rangle, which should be one or more of p, T, F and TF.

\begin{verbatim}
\prg_return_true: * \prg_return_false: *
\end{verbatim}

These “return” functions define the logical state of a conditional statement. They appear within the code for a conditional function generated by \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 \prg_return_true: or \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.

\begin{verbatim}
\prg_generate_conditional_variant:Nnn \langle name \rangle:\langle arg spec \rangle {\langle variant argument specifiers \rangle} {\langle condition specifiers \rangle}
\end{verbatim}

Defines argument-specifier variants of conditionals. This is equivalent to running \cs_generate_variant:Nn \langle conditional \rangle {\langle variant argument specifiers \rangle} on each \langle conditional \rangle described by the \langle condition specifiers \rangle. These base-form \langle conditionals \rangle are obtained from the \langle name \rangle and \langle arg spec \rangle as described for \prg_new_conditional:Npnn, and they should be defined.

### 9.2 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 (e.g., draft/final) and other times you perhaps want to use it as a predicate function in an \if predicate:w test. The problem of the primitive \if_false: and \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: \c_true_bool or \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, etc. which can then be used on both the boolean type and predicate functions.
All conditional \texttt{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{TeXhacker note:} The \texttt{bool} data type is not implemented using the \texttt{iffalse/iftrue} primitives, in contrast to \texttt{neviz}, etc., in plain \TeX. I\TeX2\varepsilon and so on. Programmers should not base use of \texttt{bool} switches on any particular expectation of the implementation.

\[
\begin{align*}
\texttt{bool_new:N} & \quad \texttt{bool_new:N \langle boolean \rangle} \\
\texttt{bool_new:c} & \quad \texttt{bool_new:c} \\
\texttt{bool_const:Nn} & \quad \texttt{bool_const:Nn \langle boolean \rangle \{\langle boolexpr \rangle\}} \\
\texttt{bool_set_false:N} & \quad \texttt{bool_set_false:N \langle boolean \rangle} \\
\texttt{bool_set_false:c} & \quad \texttt{bool_set_false:c} \\
\texttt{bool_gset_false:N} & \quad \texttt{bool_gset_false:N \langle boolean \rangle} \\
\texttt{bool_gset_false:c} & \quad \texttt{bool_gset_false:c} \\
\texttt{bool_set_true:N} & \quad \texttt{bool_set_true:N \langle boolean \rangle} \\
\texttt{bool_set_true:c} & \quad \texttt{bool_set_true:c} \\
\texttt{bool_gset_true:N} & \quad \texttt{bool_gset_true:N \langle boolean \rangle} \\
\texttt{bool_gset_true:c} & \quad \texttt{bool_gset_true:c} \\
\texttt{bool_set_eq:NN} & \quad \texttt{bool_set_eq:NN \langle boolean \rangle \langle boolean \rangle} \\
\texttt{bool_set_eq:(cN|Nc|cc)} & \quad \texttt{bool_set_eq:(cN|Nc|cc)} \\
\texttt{bool_gset_eq:NN} & \quad \texttt{bool_gset_eq:NN \langle boolean \rangle \langle boolean \rangle} \\
\texttt{bool_gset_eq:(cN|Nc|cc)} & \quad \texttt{bool_gset_eq:(cN|Nc|cc)} \\
\texttt{bool_set:Nn} & \quad \texttt{bool_set:Nn \langle boolean \rangle \{\langle boolexpr \rangle\}} \\
\texttt{bool_set:cn} & \quad \texttt{bool_set:cn} \\
\texttt{bool_gset:Nn} & \quad \texttt{bool_gset:Nn \langle boolean \rangle \{\langle boolexpr \rangle\}} \\
\texttt{bool_gset:cn} & \quad \texttt{bool_gset:cn} \\
\texttt{bool_show:N} & \quad \texttt{bool_show:N \langle boolean \rangle} \\
\texttt{bool_show:c} & \quad \texttt{bool_show:c} \\
\end{align*}
\]
\bool_show:n \bool_show:n \{\langle\text{boolean expression}\rangle\}

Displays the logical truth of the \langle\text{boolean expression}\rangle on the terminal.

\bool_log:N \bool_log:c \bool_log:n
\bool_log:N \{\langle\text{boolean}\rangle\}

Writes the logical truth of the \langle\text{boolean}\rangle in the log file.

\bool_log:n \bool_log:n
\bool_log:n \{\langle\text{boolean expression}\rangle\}

Writes the logical truth of the \langle\text{boolean expression}\rangle in the log file.

\bool_if_exist_p:N \bool_if_exist_p:c \bool_if_exist:N \bool_if_exist:c
\bool_if_exist_p:N \langle\text{boolean}\rangle
\bool_if_exist_p:N \langle\text{true code}\rangle \{\langle\text{false code}\rangle\}

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

9.2.1 Scratch booleans

\l_tmpa_bool \l_tmpb_bool
\l_tmpa_bool \l_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.

\g_tmpa_bool \g_tmpb_bool
\g_tmpa_bool \g_tmpb_bool

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.

9.3 Boolean expressions

As we have a boolean datatype and predicate functions returning boolean \langle\text{true}\rangle or \langle\text{false}\rangle values, it seems only fitting that we also provide a parser for \langle\text{boolean expressions}\rangle.

A boolean expression is an expression which given input in the form of predicate functions and boolean variables, return boolean \langle\text{true}\rangle or \langle\text{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 \}
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.

\textbf{\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 \texttt{#1} were a closing parenthesis and \texttt{\l_tmpa_bool} were \texttt{true}.

\begin{verbatim}
( \l_tmpa_bool || \token_if_eq_meaning_p:NN X #1 )
\end{verbatim}

Minimal (lazy) evaluation can be obtained using the conditionals \texttt{\bool_lazy_all:nTF}, \texttt{\bool_lazy_and:nnTF}, \texttt{\bool_lazy_any:nTF}, or \texttt{\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

\begin{verbatim}
\bool_lazy_and_p:nn
{ \bool_lazy_any_p:n
{ \int_compare_p:n { 2 = 3 } }
{ \int_compare_p:n { 4 <= 4 } }
\int_compare_p:n { 1 = \error } \% skipped
}
{ ! \int_compare_p:n { 2 = 4 } }
\end{verbatim}

the line marked with \texttt{skipped} is not expanded because the result of \texttt{\bool_lazy_any_p:n} is known once the second boolean expression is found to be logically \texttt{true}. On the other hand, the last line is expanded because its logical value is needed to determine the result of \texttt{\bool_lazy_and_p:nn}.

\begin{verbatim}
\bool_if_p:n \bool_if:nTF

\bool_if_p:n \bool_if:nTF \{ (boolean expression) \} \{ (true code) \} \{ (false code) \}
\end{verbatim}

Tests the current truth of (boolean expression), and continues expansion based on this result. The (boolean expression) 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

\bool_lazy_all_p:n \{ \{booleanexpr\} \{booleanexpr\} \cdots \{booleanexpr\} \} \{ (true code) \} \{ (false code) \}
\end{verbatim}

Implements the “And” operation on the (boolean expressions), hence is \texttt{true} if all of them are \texttt{true} and \texttt{false} if any of them is \texttt{false}. Contrarily to the infix operator &&, only the (boolean expressions) which are needed to determine the result of \texttt{\bool_lazy_all:nTF} are evaluated. See also \texttt{\bool_lazy_and:nnTF} when there are only two (boolean expressions).
\bool_lazy_and_p:nn \boolexpr{1} \boolexpr{2}
\bool_lazy_and:nnTF \boolexpr{1} \boolexpr{2} \truecode \falsecode

Implements the “And” operation between two boolean expressions, hence is true if both are true. Contrarily to the infix operator \&\&, the \boolexpr{2} 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 boolean expressions.

\bool_lazy_or_p:nn \boolexpr{1} \boolexpr{2}
\bool_lazy_or:nnTF \boolexpr{1} \boolexpr{2} \truecode \falsecode

Implements the “Or” operation between two boolean expressions, hence is true if either one is true. Contrarily to the infix operator ||, the \boolexpr{2} is only evaluated if it is needed to determine the result of \bool_lazy_or:nnTF. See also \bool_lazy_or:nnTF when there are only two boolean expressions.

\bool_not_p:n \boolexpr{1}

Function version of !(\boolexpr{1}) within a boolean expression.

\bool_xor_p:nn \boolexpr{1} \boolexpr{2}
\bool_xor:nnTF \boolexpr{1} \boolexpr{2} \truecode \falsecode

Implements an “exclusive or” operation between two boolean expressions. There is no infix operation for this logical operation.

9.4 Logical loops

Loops using either boolean expressions or stored boolean values.

\bool_do_until:Nn \boolexpr{1} \code
\bool_do_until:cn \boolexpr{1} \code

Places the \code in the input stream for \TeX to process, and then checks the logical value of the \boolexpr{1}. If it is false then the \code is inserted into the input stream again and the process loops until the \boolexpr{1} is true.

\bool_do_while:Nn \boolexpr{1} \code
\bool_do_while:cn \boolexpr{1} \code

Places the \code in the input stream for \TeX to process, and then checks the logical value of the \boolexpr{1}. If it is true then the \code is inserted into the input stream again and the process loops until the \boolexpr{1} is false.
\bool_until_do:Nn \bool_until_do:cn

This function first checks the logical value of the \texttt{boolean}. If it is \texttt{false} the \texttt{code} is placed in the input stream and expanded. After the completion of the \texttt{code} the truth of the \texttt{boolean} is re-evaluated. The process then loops until the \texttt{boolean} is \texttt{true}.

\bool_while_do:Nn \bool_while_do:cn

This function first checks the logical value of the \texttt{boolean}. If it is \texttt{true} the \texttt{code} is placed in the input stream and expanded. After the completion of the \texttt{code} the truth of the \texttt{boolean} is re-evaluated. The process then loops until the \texttt{boolean} is \texttt{false}.

\bool_do_until:nn \bool_do_until:cn

Places the \texttt{code} in the input stream for \LaTeX{} to process, and then checks the logical value of the \texttt{boolean expression} as described for \texttt{\bool_if:nTF}. If it is \texttt{false} then the \texttt{code} is inserted into the input stream again and the process loops until the \texttt{boolean expression} evaluates to \texttt{true}.

\bool_do_while:nn \bool_do_while:cn

Places the \texttt{code} in the input stream for \LaTeX{} to process, and then checks the logical value of the \texttt{boolean expression} as described for \texttt{\bool_if:nTF}. If it is \texttt{true} then the \texttt{code} is inserted into the input stream again and the process loops until the \texttt{boolean expression} evaluates to \texttt{false}.

\bool_until_do:nn \bool_until_do:cn

This function first checks the logical value of the \texttt{boolean expression} (as described for \texttt{\bool_if:nTF}). If it is \texttt{false} the \texttt{code} is placed in the input stream and expanded. After the completion of the \texttt{code} the truth of the \texttt{boolean expression} is re-evaluated. The process then loops until the \texttt{boolean expression} is \texttt{true}.

\bool_while_do:nn \bool_while_do:cn

This function first checks the logical value of the \texttt{boolean expression} (as described for \texttt{\bool_if:nTF}). If it is \texttt{true} the \texttt{code} is placed in the input stream and expanded. After the completion of the \texttt{code} the truth of the \texttt{boolean expression} is re-evaluated. The process then loops until the \texttt{boolean expression} is \texttt{false}.

\section{Producing multiple copies}

\prg_replicate:nn \prg_replicate:cn

Evaluates the \texttt{integer expression} (which should be zero or positive) and creates the resulting number of copies of the \texttt{tokens}. The function is both expandable and safe for nesting. It yields its result after two expansion steps.
9.6 Detecting \TeX{}’s mode

\ifx\mode_if_horizontal_p\undefined\else\fi
\if\mode_if_horizontal:TF \fi
\langle \true\rangle
\else\fi
\langle \false\rangle
Detects if \TeX{} is currently in horizontal mode.

\ifx\mode_if_inner_p\undefined\else\fi
\if\mode_if_inner:TF \fi
\langle \true\rangle
\else\fi
\langle \false\rangle
Detects if \TeX{} is currently in inner mode.

\ifx\mode_if_math_p\undefined\else\fi
\if\mode_if_math:TF \fi
\langle \true\rangle
\else\fi
\langle \false\rangle
Detects if \TeX{} is currently in maths mode.

\ifx\mode_if_vertical_p\undefined\else\fi
\if\mode_if_vertical:TF \fi
\langle \true\rangle
\else\fi
\langle \false\rangle
Detects if \TeX{} is currently in vertical mode.

9.7 Primitive conditionals

\if\if_predicate:w\fi
\langle \predicate\rangle \langle \true\rangle \else:\langle \false\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 \if_bool:N.)

\if\if_bool:N\fi
\langle \boolean\rangle \langle \true\rangle \else:\langle \false\rangle \fi:
This function takes a boolean variable and branches according to the result.

9.8 Nestable recursions and mappings

There are a number of places where recursion or mapping constructs are used in 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.

\if\prg_break_point:Nn\fi
\langle \type\rangle\_map\break: \langle \code\rangle
Used to mark the end of a recursion or mapping: the functions \langle \type\rangle\_map\break: and \langle \type\rangle\_map\break:n use this to break out of the loop (see \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: \prg_break_point:Nn is functionally-equivalent in these cases to \use_ii:nn.
\prg_map_break:Nn \( \text{type} \)_map_break: \{ \langle \text{user code} \rangle \} \\
\prg_break_point:Nn \( \text{type} \)_map_break: \{ \langle \text{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 \( \text{type} \)_-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, \( \text{type} \)_map_break: and \( \text{type} \)_-map_break:n are defined as \prg_map_break:Nn \( \text{type} \)_map_break: {} and the same with {} omitted.

9.8.1 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.

\prg_break_point: * \\
\prg_break: * \\
\prg_break:n \{ \langle \text{code} \rangle \} \ldots \prg_break_point: 

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 \text{code} \rangle \} \ldots \prg_break_point: 

Breaks a recursion which has no \langle ending code \rangle and which is not a user-breakable mapping (see for instance \prop_get:Nn), and inserts the \langle code \rangle in the input stream.

9.9 Internal programming functions

\group_align_safe_begin: * \\
\group_align_safe_end: * \\
\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.
Chapter 10

The \texttt{l3sys} package:
System/runtime functions

10.1 The name of the job

\begin{verbatim}
\c_sys_jobname_str
\end{verbatim}

Constant that gets the “job name” assigned when \TeX{} starts.

\TeX{}hackers 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.

10.2 Date and time

\begin{verbatim}
\c_sys_minute_int
\c_sys_hour_int
\c_sys_day_int
\c_sys_month_int
\c_sys_year_int
\end{verbatim}

The date and time at which the current job was started: these are all reported as integers.

\TeX{}hackers 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.
10.3 Engine

Conditionals which allow engine-specific code to be used. The names follow naturally from those of the engine binaries: note that the (u)p\texttt{p}tex tests are for \(\varepsilon\text{-p}\TeX\) and \(\varepsilon\text{-up}\TeX\) as expl3 requires the \(\varepsilon\text{-}\TeX\) extensions. Each conditional is true for \textit{exactly one} supported engine. In particular, \texttt{sys_if_engine_ptex_p:} is true for \(\varepsilon\text{-p}\TeX\) but false for \(\varepsilon\text{-up}\TeX\).

\begin{verbatim}
\sys_if_engine_luatex_p:  *
\sys_if_engine_luatex:TF *
\sys_if_engine_pdfTeX_p:  *
\sys_if_engine_pdfTeX:TF *
\sys_if_engine_ptex_p:    *
\sys_if_engine_ptex:TF *
\sys_if_engine_upTeX_p:   *
\sys_if_engine_upTeX:TF *
\sys_if_engine_xetex_p:   *
\sys_if_engine_xetex:TF *
\end{verbatim}

New: 2015-09-07

The current engine given as a lower case string: one of \texttt{luatex}, \texttt{pdftex}, \texttt{ptex}, \texttt{uptex} or \texttt{xetex}.

\begin{verbatim}
\c_sys_engine_str
\end{verbatim}

New: 2015-09-19

The name of the standard executable for the current \TeX\ engine given as a lower case string: one of \texttt{luatex}, \texttt{luahbtex}, \texttt{pdftex}, \texttt{eptex}, \texttt{euptex} or \texttt{xetex}.

\begin{verbatim}
\c_sys_engine_exec_str
\end{verbatim}

New: 2020-08-20

The name of the preloaded format for the current \TeX\ run given as a lower case string: one of \texttt{lualatex} (or \texttt{dvilualatex}), \texttt{pdflatex} (or \texttt{latex}), \texttt{platex}, \texttt{uplatex} or \texttt{xelatex} for \(\varepsilon\text{-}\TeX\), similar names for plain \TeX\ (except \texttt{pdf}\TeX\ in DVI mode yields \texttt{etex}), and \texttt{cont-en} for Con\TeX\ (i.e. the \texttt{fmtname}).

\begin{verbatim}
\sys_timer:
\end{verbatim}

New: 2020-09-24

Expands to the current value of the engine’s timer clock, a non-negative integer. This function is only defined for engines with timer support. This command measures not just CPU time but real time (including time waiting for user input). The unit are scaled seconds (\(2^{-16}\) seconds).

10.4 Output format

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.

\begin{verbatim}
\sys_if_output_dvi_p:  *
\sys_if_output_dvi:TF *
\sys_if_output_pdf_p:  *
\sys_if_output_pdf:TF *
\end{verbatim}

New: 2015-09-19

The current output mode given as a lower case string: one of \texttt{dvi} or \texttt{pdf}.

\begin{verbatim}
\c_sys_output_str
\end{verbatim}

New: 2015-09-19

71
10.5 Platform

\sys_if_platform_unix_p: \* \sys_if_platform_unix:TF \{(true code)\} \{(false code)\}
\sys_if_platform_unix:TF \*
\sys_if_platform_windows_p: \*
\sys_if_platform_windows:TF \*

New: 2018-07-27

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.

10.6 Random numbers

\sys_rand_seed: \*

\sys_rand_seed:

New: 2017-05-27

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 \{(intexpr)\}

New: 2017-05-27

Globally sets the seed for the engine’s pseudo-random number generator to the \{(integer expression)\}. 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 \([0, 2^{28} - 1]\).

10.7 Access to the shell

\sys_get_shell:nnN \{(shell command)\} \{(setup)\} \{(tl var)\} \{(true code)\} \{(false code)\}
\sys_get_shell:nnNTF \{(shell command)\} \{(setup)\} \{(tl var)\} \{(true code)\} \{(false code)\}

New: 2019-09-20

Defines \(tl\) to the text returned by the \{(shell command)\}. The \{(shell command)\} is converted to a string using \tl_to_str:n. Category codes may need to be set appropriately via the \{(setup)\} argument, which is run just before running the \{(shell command)\} (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 \{(shell command)\}. The \sys_get_shell:nnNTF conditional returns true if the shell is available and no quote is detected, and false otherwise.
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_p:
\sys_if_shell:TF  \sys_if_shell:TF \{(true code)\} \{(false code)\}

Performs a check for whether shell escape is enabled. This returns true if either of restricted or unrestricted shell escape is enabled.

\sys_if_shell_unrestricted_p:  \sys_if_shell_unrestricted_p:
\sys_if_shell_unrestricted:TF  \sys_if_shell_unrestricted:TF \{(true code)\} \{(false code)\}

Performs a check for whether unrestricted shell escape is enabled.

\sys_if_shell_restricted_p:  \sys_if_shell_restricted_p:
\sys_if_shell_restricted:TF  \sys_if_shell_restricted:TF \{(true code)\} \{(false code)\}

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:.

\sys_shell_now:n  \sys_shell_now:n \{(tokens)\}
\sys_shell_now:x  \sys_shell_now:x

Execute \(\text{tokens}\) through shell escape immediately.

\sys_shell_shipout:n  \sys_shell_shipout:n \{(tokens)\}
\sys_shell_shipout:x  \sys_shell_shipout:x

Execute \(\text{tokens}\) through shell escape at shipout.

### 10.8 Loading configuration data

\sys_load_backend:n  \sys_load_backend:n \{(backend)\}
\sys_load_backend:n

 Loads the additional configuration file needed for backend support. If the \(\text{backend}\) 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.
Load the additional configuration files for debugging support and rolling back depreca-
tions, respectively.

10.8.1 Final settings

Finalises all system-dependent functionality: required before loading a backend.
Chapter 11

The \texttt{l3msg} package

Messages

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 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.

11.1 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, \texttt{\textbackslash\textbackslash} may be used to force a new line and \texttt{\textbackslash u} forces an explicit space. Additionally, \texttt{\textbackslash!, \textbackslash#, \textbackslash%, \textbackslash~} and \texttt{\textbackslash -} can be used to produce the corresponding character.

Messages may be subdivided \texttt{by one level} using the \texttt{/} character. This is used within the message filtering system to allow for example the \texttt{\LaTeX} kernel messages to belong to the module \texttt{\LaTeX} while still being filterable at a more granular level. Thus for example

\texttt{\msg\_new:nnnn \{ mymodule \} \{ submodule / message \} ...}

will allow to filter out specifically messages from the \texttt{submodule}.
\msg_new:nnnn \msg_new:nnnn

Updated: 2011-08-16

\msg_new:nnnn \msg_new:nnnn
\msg_set:nnnn \msg_set:nnn
\msg_set:nnnn \msg_set:nnnn
\msg_gset:nnnn \msg_gset:nnnn

\msg_if_exist_p:nn \msg_if_exist_p:nn
\msg_if_exist:nnTF \msg_if_exist:nnTF

\msg_module_name:n \msg_module_name:n
\msg_module_type:n \msg_module_type:n

\msg_line_context:

11.2 Customizable information for message modules

\msg_module_name:n \msg_module_name:n

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

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.

\g_msg_module_name_prop

Provides a mapping between the module name used for messages, and that for documentation. For example, \LaTeX3 core messages are stored in the reserved \LaTeX tree, but are printed as \LaTeX3.

\g_msg_module_type_prop

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.

11.3 Contextual information for messages

\msg_line_context:

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 on line.
\msg_line_number: * \msg_line_number:
Prints the current line number when a message is given.

\msg_fatal_text:n * \msg_fatal_text:n \{\langle module\rangle}\}
Produces the standard text

Fatal Package \langle module \rangle Error
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.

\msg_critical_text:n * \msg_critical_text:n \{\langle module\rangle}\}
Produces the standard text

Critical Package \langle module \rangle Error
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.

\msg_error_text:n * \msg_error_text:n \{\langle module\rangle}\}
Produces the standard text

Package \langle module \rangle Error
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.

\msg_warning_text:n * \msg_warning_text:n \{\langle module\rangle}\}
Produces the standard text

Package \langle module \rangle Warning
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_info_text:n * \msg_info_text: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.
11.4 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. Note that this expansion takes place with the standard definitions in effect, which means that shorthands such as \- or \\ are not available; instead one should use \iow_char:N \~ and \iow_newline:, respectively. The following message classes exist:

- **fatal**, ending the \TeX{} run;
- **critical**, ending the file being input;
- **error**, interrupting the \TeX{} run without ending it;
- **warning**, written to terminal and log file, for important messages that may require corrections by the user;
- **note** (less common than **info**) for important information messages written to the terminal and log file;
- **info** for normal information messages written to the log file only;
- **term** and **log** for un-decorated messages written to the terminal and log file, or to the log file only;
- **none** for suppressed messages.

\msg_fatal:nnnnn \{\langle module\rangle\} \{\langle message\rangle\} \{\langle arg one\rangle\} \{\langle arg two\rangle\} \{\langle arg three\rangle\} \{\langle arg four\rangle\}

Issues \langle module\rangle error \langle message\rangle, passing \langle arg one\rangle to \langle arg four\rangle to the text-creating functions. After issuing a fatal error the \TeX{} run halts. No PDF file will be produced in this case (DVI mode runs may produce a truncated DVI file).
\msg_critical:nnnnnn \{\text{module}\} \{\text{message}\} \{\text{arg one}\} \{\text{arg two}\} \{\text{arg three}\} \{\text{arg four}\}

Issues \textit{(module)} error \textit{(message)}, passing \textit{(arg one)} to \textit{(arg four)} to the text-creating functions. After issuing a critical error, \TeX stops reading the current input file. This may halt the \TeX run (if the current file is the main file) or may abort reading a sub-file.

\TeXhackers\textit{note}: The \TeX \texttt{\textbackslash endinput} primitive is used to exit the file. In particular, the rest of the current line remains in the input stream.

---

\msg_error:nnnnnn \{\text{module}\} \{\text{message}\} \{\text{arg one}\} \{\text{arg two}\} \{\text{arg three}\} \{\text{arg four}\}

Issues \textit{(module)} error \textit{(message)}, passing \textit{(arg one)} to \textit{(arg four)} to the text-creating functions. The error interrupts processing and issues the text at the terminal. After user input, the run continues.

---

\msg_warning:nnnnnn \{\text{module}\} \{\text{message}\} \{\text{arg one}\} \{\text{arg two}\} \{\text{arg three}\} \{\text{arg four}\}

Issues \textit{(module)} warning \textit{(message)}, passing \textit{(arg one)} to \textit{(arg four)} to the text-creating functions. The warning text is added to the log file and the terminal, but the \TeX run is not interrupted.
Issues ⟨module⟩ information ⟨message⟩, passing ⟨arg one⟩ to ⟨arg four⟩ to the text-creating functions. For the more common \msg_info:nnnnnn, the information text is added to the log file only, while \msg_note:nnnnnn adds the info text to both the log file and the terminal. The \TeX run is not interrupted.

\msg_note:nnnnnn ⟨module⟩} ⟨message⟩} ⟨arg one⟩} ⟨arg two⟩} ⟨arg three⟩} ⟨arg four⟩
\msg_info:nnnnnn ⟨module⟩} ⟨message⟩} ⟨arg one⟩} ⟨arg two⟩} ⟨arg three⟩} ⟨arg four⟩

Issues ⟨module⟩ information ⟨message⟩, passing ⟨arg one⟩ to ⟨arg four⟩ to the text-creating functions. The output is briefer than \msg_info:nnnnnn, omitting for instance the module name. It is added to the log file by \msg_log:nnnnnn while \msg_term:nnnnnn also prints it on the terminal.

\msg_term:nnnnnn ⟨module⟩} ⟨message⟩} ⟨arg one⟩} ⟨arg two⟩} ⟨arg three⟩} ⟨arg four⟩
\msg_log:nnnnnn ⟨module⟩} ⟨message⟩} ⟨arg one⟩} ⟨arg two⟩} ⟨arg three⟩} ⟨arg four⟩

Does nothing: used as a message class to prevent any output at all (see the discussion of message redirection).

\msg_none:nnnnnn ⟨module⟩} ⟨message⟩} ⟨arg one⟩} ⟨arg two⟩} ⟨arg three⟩} ⟨arg four⟩
11.4.1 Messages for showing material

\msg_show:nnnnnn {⟨module⟩} {⟨message⟩} {⟨arg one⟩} {⟨arg two⟩} {⟨arg three⟩} {⟨arg four⟩}

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 \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 \> at the start of a line, an additional line \> will be put at the end. In addition, a final period is added if not present.

11.4.2 Expandable error messages

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_expandable_error:nnnnnn * \msg_expandable_error:nnffff * \msg_expandable_error:nnnnn * \msg_expandable_error:nnfff * \msg_expandable_error:nnnn * \msg_expandable_error:nnff * \msg_expandable_error:nnn * \msg_expandable_error:nf * \msg_expandable_error:nn *

Issues an “Undefined error” message from \TeX itself using the undefined control sequence \.:error then prints “! ⟨module⟩: ”⟨error message⟩, which should be short. With default settings, anything beyond approximately 50 characters long (or bytes in some engines) is cropped. A leading space might be removed as well.

11.5 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 }

81
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 $A \rightarrow B$, $B \rightarrow C$ and $C \rightarrow A$ in this order, then the $A \rightarrow B$ redirection is
cancelled.

Updated: 2012-04-27

```
\msg_redirect_class:nn {⟨class one⟩} {⟨class two⟩}
```
Changes the behaviour of messages of ⟨class one⟩ so that they are processed using the
code for those of ⟨class two⟩. Each ⟨class⟩ can be one of fatal, critical, error, warning, note, info, term, log, none.

Updated: 2012-04-27

```
\msg_redirect_module:nnn { ⟨module⟩ } { ⟨class one⟩ } { ⟨class two⟩ }
```
Redirects message of ⟨class one⟩ for ⟨module⟩ to act as though they were from ⟨class
two⟩. Messages of ⟨class one⟩ from sources other than ⟨module⟩ are not affected by this
redirection. This function can be used to make some messages “silent” by default. For
example, all of the warning messages of ⟨module⟩ could be turned off with:

```
\msg_redirect_module:nnn { module } { warning } { none }
```

Updated: 2012-04-27

```
\msg_redirect_name:nnn { ⟨module⟩ } { ⟨message⟩ } { ⟨class⟩ }
```
Redirects a specific ⟨message⟩ from a specific ⟨module⟩ to act as a member of ⟨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 }
```
Chapter 12

The \texttt{l3file} 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{\textbackslash file\_...}, while others are used to work with files on a line by line basis and have prefix \texttt{\textbackslash ior\_...} (reading) or \texttt{\textbackslash 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 \textit{not} expanded, allowing the direct use of these in file names. Quote tokens (\texttt{	extquotedbl}) 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.

12.1 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 \LaTeX{}3. 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.
Globally reserves the name of the \( \text{stream} \), either for reading or for writing as appropriate. The \( \text{stream} \) is not opened until the appropriate \texttt{\ior\_open:Nn} function is used. Attempting to use a \( \text{stream} \) which has not been opened is an error, and the \( \text{stream} \) will behave as the corresponding \texttt{c\_term\_...}.

\texttt{\ior\_open:Nn} \{\langle file name\rangle\}

Opens \( \langle file name\rangle \) for reading using \( \langle stream\rangle \) as the control sequence for file 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 \( \langle file name\rangle \) until a \texttt{\ior\_close:N} instruction is given or the \TeX run ends. If the file is not found, an error is raised.

\texttt{\ior\_open:NnTF} \{\langle file name\rangle\} \{\langle true code\rangle\} \{\langle false code\rangle\}

Opens \( \langle file name\rangle \) for reading using \( \langle stream\rangle \) as the control sequence for file 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 \( \langle file name\rangle \) until a \texttt{\ior\_close:N} instruction is given or the \TeX run ends. The \( \langle true code\rangle \) is then inserted into the input stream. If the file is not found, no error is raised and the \( \langle false code\rangle \) is inserted into the input stream.

\texttt{\iow\_open:Nn} \{\langle file name\rangle\}

Opens \( \langle file name\rangle \) for writing using \( \langle stream\rangle \) as the control sequence for file 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 \( \langle file name\rangle \) until a \texttt{\iow\_close:N} instruction is given or the \TeX run ends. Opening a file for writing clears any existing content in the file (i.e. writing is not additive).

\texttt{\ior\_close:N} \langle stream\rangle

Closes the \( \langle stream\rangle \). Streams should always be closed when they are finished with as this ensures that they remain available to other programmers.

\texttt{\ior\_show:N} \langle stream\rangle

Display (to the terminal or log file) the file name associated to the \( \langle \text{read or write} \rangle \) \( \langle \text{stream} \rangle \).
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.

12.1.1 Reading from files

Reading from files and reading from the terminal are separate processes in expl3. The functions \ior_get:NN and \ior_str_get:NN, and their branching equivalents, are designed to work with files.

\ior_get:NN \ior_get:NNTF

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 \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 \texttt{a b c \par}. Any blank line is converted to the token \texttt{\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 \texttt{\par} tokens. In the non-branching version, where the \langle stream \rangle is not open the \langle tl var \rangle is set to \texttt{\q_no_value}.

\textbf{\TeXhackers note:} This protected macro is a wrapper around the \TeX primitive \texttt{\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 \texttt{\endlinechar}, omitted if \texttt{\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.
\ior_str_get:NN \ior_str_get:NN \langle stream \rangle \langle token list variable \rangle
\ior_str_get:NN \ior_str_get:NN \langle stream \rangle \langle token list variable \rangle \langle true code \rangle \langle false code \rangle

Function that reads one line from the file input \langle stream \rangle and stores the result locally in the \langle token list \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 \texttt{a b c} with the letters \texttt{a}, \texttt{b}, and \texttt{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.

\textbf{TeXHackers note:} This protected macro is a wrapper around the $\texttt{\varepsilon}$-\TeX primitive \texttt{\ 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, \emph{i.e.} 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 \langle stream \rangle \{ \langle inline function \rangle \}

Applies the \langle inline function \rangle to each set of \langle lines \rangle 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 \langle inline function \rangle should consist of code which receives the \langle line \rangle as \#1.

\ior_str_map_inline:Nn \ior_str_map_inline:Nn \langle stream \rangle \{ \langle inline function \rangle \}

Applies the \langle inline function \rangle to every \langle line \rangle in the \langle stream \rangle. 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). The \langle inline function \rangle should consist of code which receives the \langle line \rangle as \#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 \langle stream \rangle \langle tl var \rangle \{ \langle code \rangle \}

For each set of \langle lines \rangle obtained by calling \ior_get:NN until reaching the end of the file, stores the \langle lines \rangle in the \langle tl var \rangle then 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 set of \langle lines \rangle, or its original value if the \langle stream \rangle is empty. \TeX ignores any trailing new-line marker from the file it reads. This function is typically faster than \ior_map_inline:Nn.
\ior_str_map_variable:NNn \ior_str_map_variable:NNn \stream \variable \{\code\}

For each \line in the \stream, stores the \line in the \variable then applies the \code. The material is read from the \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 \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 \line, or its original value if the \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 \stream have been processed. 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:}
{ % Do something useful
}
}

Use outside of a \ior_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.

\ior_map_break:n \ior_map_break:n \{\code\}

Used to terminate a \ior_map... function before all lines in the \stream have been processed, inserting the \code 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.

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

12.1.2 Writing to files

\texttt{\texttt{\textbackslash iow\_now:Nn}} \hspace{1cm} \texttt{\textbackslash iow\_now:(N\texttt{\textbackslash x|cn|cx})}

This function writes \langle tokens\rangle to the specified \langle stream\rangle immediately (i.e. the write operation is called on expansion of \texttt{\textbackslash iow\_now:Nn}).

\texttt{\texttt{\textbackslash iow\_log:n}} \hspace{1cm} \texttt{\textbackslash iow\_log:x}

This function writes the given \langle tokens\rangle to the log (transcript) file immediately: it is a dedicated version of \texttt{\textbackslash iow\_now:Nn}.

\texttt{\texttt{\textbackslash iow\_term:n}} \hspace{1cm} \texttt{\textbackslash iow\_term:x}

This function writes the given \langle tokens\rangle to the terminal file immediately: it is a dedicated version of \texttt{\textbackslash iow\_now:Nn}.

\texttt{\texttt{\textbackslash iow\_shipout:Nn}} \hspace{1cm} \texttt{\textbackslash iow\_shipout:(N\texttt{\textbackslash x|cn|cx})}

This functions writes \langle tokens\rangle to the specified \langle stream\rangle when the current page is finalised (i.e. at shipout). The x-type variants expand the \langle tokens\rangle at the point where the function is used but not when the resulting tokens are written to the \langle stream\rangle (cf. \texttt{\textbackslash iow\_shipout\_x:Nn}).

\textbf{TEXhackers note:} When using expl3 with a format other than \LaTeX{}, new line characters inserted using \texttt{\textbackslash iow\_newline:} or using the line-wrapping code \texttt{\textbackslash iow\_wrap:mmNN} are not recognized in the argument of \texttt{\textbackslash iow\_shipout:Nn}. This may lead to the insertion of additional unwanted line-breaks.

\texttt{\texttt{\textbackslash iow\_shipout\_x:Nn}} \hspace{1cm} \texttt{\textbackslash iow\_shipout\_x:(N\texttt{\textbackslash x|cn|cx})}

This functions writes \langle tokens\rangle to the specified \langle stream\rangle when the current page is finalised (i.e. at shipout). The \langle tokens\rangle 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{TEXhackers note:} This is a wrapper around the \LaTeX{} primitive \texttt{\textbackslash write}. When using expl3 with a format other than \LaTeX{}, new line characters inserted using \texttt{\textbackslash iow\_newline:} or using the line-wrapping code \texttt{\textbackslash iow\_wrap:mmNN} are not recognized in the argument of \texttt{\textbackslash iow\_shipout:Nn}. This may lead to the insertion of additional unwanted line-breaks.
\iow_char:N \iow_char:N \(\text{char}\)

Inserts \texttt{\(\text{char}\)} into the output stream. Useful when trying to write difficult characters such as \texttt{\%}, \texttt{\{}, \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:

Function to add a new line within the \texttt{\textit{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{\TeX} hackers note: When using \texttt{expl3} with a format other than \texttt{\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.
12.1.3 Wrapping lines in output

\iow_wrap:nnnN \iow_wrap:nnnN \{\text\} \{\run-on text\} \{\set up\} \{\function\}

This function wraps the \text to a fixed number of characters per line. At the start of each line which is wrapped, the \run-on text is inserted. The line character count targeted is the value of \l_iow_line_count_int minus the number of characters in the \run-on text 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 \text and \run-on text 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 \text (not the \run-on text).

Additional functions may be added to the wrapping by using the \set up, which is executed before the wrapping takes place: this may include overriding the substitutions listed.

Any expandable material in the \text 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 \function, which is typically a wrapper around a write operation. The output of \iow_wrap:nnnN (i.e. the argument passed to the \function) 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.

\TeX hackers note: Internally, \iow_wrap:nnnN carries out an x-type expansion on the \text 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 \text.

\iow_indent:n \iow_indent:n \{\text\}

In the first argument of \iow_wrap:nnnN (for instance in messages), indents \text by four spaces. This function does not cause a line break, and only affects lines which start within the scope of the \text. In case the indented \text should appear on separate lines from the surrounding text, use \ or to force line breaks.

\l_iow_line_count_int \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.
12.1.4 Constant input–output streams, and variables

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

New: 2017-12-11

Constant output streams for writing to the log and to the terminal (plus the log), respectively.

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

New: 2017-12-11

12.1.5 Primitive conditionals

Tests if the \langle stream\rangle returns “end of file”, which is true for non-existent files. The \texttt{\else:} branch is optional.

\texttt{\TeXhackers note:} This is the \TeX primitive \texttt{\ifeof}.

12.2 File operation functions

Contain the directory, name and extension of the current file. The directory is empty if the file was loaded without an explicit path (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{\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.

**\TeX{}hackers note:** When working as a package in \LaTeX{} 2\epsilon, expl3 will automatically append the current \texttt{\input@path} to the set of values from \texttt{\l_file_search_path_seq}.

\begin{verbatim}
\file_if_exist:nTF \{\langle file name\rangle\} \{\langle true code\rangle\} \{\langle false code\rangle\}
\end{verbatim}

Searches for \texttt{\langle file name\rangle} using the current \TeX{} search path and the additional paths controlled by \texttt{\l_file_search_path_seq}.

\begin{verbatim}
\file_get:nnN \{\langle filename\rangle\} \{\langle setup\rangle\} \langle tl\rangle
\end{verbatim}

\begin{verbatim}
\file_get:nnNTF \{\langle filename\rangle\} \{\langle setup\rangle\} \langle tl\rangle \langle true code\rangle \langle false code\rangle
\end{verbatim}

Defines \texttt{\langle tl\rangle} to the contents of \texttt{\langle filename\rangle}. Category codes may need to be set appropriately via the \texttt{\langle setup\rangle} argument. The non-branching version sets the \texttt{\langle tl\rangle} to \texttt{\q_no_value} if the file is not found. The branching version runs the \texttt{\langle true code\rangle} after the assignment to \texttt{\langle tl\rangle} if the file is found, and \texttt{\langle false code\rangle} otherwise.

\begin{verbatim}
\file_get_full_name:nN \{\langle file name\rangle\}
\end{verbatim}

\begin{verbatim}
\file_get_full_name:VN \{\langle file name\rangle\}
\end{verbatim}

\begin{verbatim}
\file_get_full_name:nNTF \{\langle file name\rangle\}
\end{verbatim}

\begin{verbatim}
\file_get_full_name:VNTF \{\langle file name\rangle\}
\end{verbatim}

Searches for \texttt{\langle file name\rangle} in the path as detailed for \texttt{\file_if_exist:nTF}, and if found sets the \texttt{\langle tl var\rangle} the fully-qualified name of the file, \textit{i.e.} the path and file name. This includes an extension \texttt{.tex} when the given \texttt{\langle file name\rangle} has no extension but the file found has that extension. In the non-branching version, the \texttt{\langle tl var\rangle} will be set to \texttt{\q_no_value} in the case that the file does not exist.

\begin{verbatim}
\file_full_name:n \{\langle file name\rangle\}
\end{verbatim}

\begin{verbatim}
\file_full_name:V \{\langle file name\rangle\}
\end{verbatim}

Searches for \texttt{\langle file name\rangle} in the path as detailed for \texttt{\file_if_exist:nTF}, and if found leaves the fully-qualified name of the file, \textit{i.e.} the path and file name, in the input stream. This includes an extension \texttt{.tex} when the given \texttt{\langle file name\rangle} 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 \langle full name \rangle and splits it into three parts, each of which is returned by setting the appropriate local string variable:

- The \langle dir \rangle: everything up to the last / (path separator) in the \langle file path \rangle. As with system PATH variables and related functions, the \langle dir \rangle does not include the trailing / unless it points to the root directory. If there is no path (only a file name), \langle dir \rangle is empty.

- The \langle name \rangle: everything after the last / up to the last ., where both of those characters are optional. The \langle name \rangle may contain multiple . characters. It is empty if \langle full name \rangle consists only of a directory name.

- The \langle ext \rangle: everything after the last . (including the dot). The \langle ext \rangle is empty if there is no . after the last /.

Before parsing, the \langle full name \rangle is expanded until only non-expandable tokens remain, except that active characters are also not expanded. Quotes (") are invalid in file names and are discarded from the input.
\file_mdfive_hash:n \{⟨file name⟩\}

Searches for ⟨file name⟩ using the current \TeX{} search path and the additional paths controlled by \l_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.

\file_get_mdfive_hash:n ⟨⟨tl var⟩⟩

Sets the ⟨⟨tl var⟩⟩ to the result of applying \file_mdfive_hash:n to the ⟨⟨file⟩⟩. If the file is not found, the ⟨⟨tl var⟩⟩ will be set to \q_no_value.

\file_size:n \{⟨file name⟩\}

Searches for ⟨file name⟩ using the current \TeX{} search path and the additional paths controlled by \l_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_get_size:n ⟨⟨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 \XeTeX.

\file_timestamp:n \{⟨file name⟩\}

Searches for ⟨file name⟩ using the current \TeX{} search path and the additional paths controlled by \l_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 \XeTeX.

\file_get_timestamp:n ⟨⟨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 \XeTeX.
\file_compare_timestamp\nNn \{\text{file-1}\} \{\text{comparator}\} \{\text{file-2}\} \{\text{true code}\} \{\text{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 that 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 \texttt{Xe\LaTeX}.

\file_input\n\{\text{name}\}

Searches for (file name) in the path as detailed for \file_if_exist\nTF, and if found reads in the file as additional \texttt{EP\TeX} 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\{\text{name}\}
\file_if_exist_input\nF \{\text{name}\} \{\text{false code}\}

Searches for (file name) using the current \texttt{EP\TeX} search path and the additional paths controlled by \file_path_include\n. If found then reads in the file as additional \texttt{EP\TeX} 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:

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\n should be preferred.

\textbf{\texttt{EP\TeX}hackers note:} This function must be used on a line on its own: \texttt{EP\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 \texttt{EP\TeX} 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.
Chapter 13

The \texttt{luatex} package: \LaTeXeX-specific functions

The \LaTeXeX 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 \texttt{pdf\TeX}, \texttt{d\TeX}, \texttt{up\TeX}, or \texttt{Xe\TeX} these raise an error: use \texttt{sys_if_engine\_lua\texttt{\textunderscore t}\_\texttt{\textunderscore X}} to avoid this. Details on using Lua with the \texttt{Lua\texttt{\textunderscore T}\texttt{\textunderscore X}} engine are given in the \texttt{Lua\texttt{\textunderscore T}\texttt{\textunderscore X}} manual.

### 13.1 Breaking out to Lua

\verb|\lua_now:n {⟨token list⟩}|  
\verb|\lua_now:e |

The ⟨token list⟩ 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 ⟨Lua input⟩ is passed to the Lua interpreter for processing. Each \verb|\lua_now:n| block is treated by Lua as a separate chunk. The Lua interpreter executes the ⟨Lua input⟩ immediately, and in an expandable manner.

\TeXhackers note: \verb|\lua_now:e| is a macro wrapper around \verb|\directlua|: when \texttt{Lua\texttt{\textunderscore T}\texttt{\textunderscore X}} is in use two expansions are required to yield the result of the Lua code.

\verb|\lua_shipout:n {⟨token list⟩}|  
\verb|\lua_shipout:n|  
\verb|\lua_shipout_e:n |

The ⟨token list⟩ 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 ⟨Lua input⟩ is passed to the Lua interpreter when the current page is finalised (\texttt{i.e.} at shipout). Each \verb|\lua_shipout:n| block is treated by Lua as a separate chunk. The Lua interpreter will execute the ⟨Lua input⟩ during the page-building routine: no \TeX expansion of the ⟨Lua input⟩ will occur at this stage.

In the case of the \verb|\lua_shipout_e:n| version the input is fully expanded by \TeX in an \texttt{e}-type manner during the shipout operation.

\TeXhackers note: At a \TeX level, the ⟨Lua input⟩ is stored as a “whatsit.”
\texttt{\textbackslash lua\_escape:n} \{\textit{token list}\}

Converts the \{\textit{token list}\} 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{\textbackslash n} and \texttt{\textbackslash r}, respectively.

\textbf{\textTeX{}hackers note:} \texttt{\textbackslash lua\_escape:e} is a macro wrapper around \texttt{\textbackslash lua\_escapestring}: when \textTeX{}\LaTeX{} is in use two expansions are required to yield the result of the Lua code.

\section*{13.2 Lua interfaces}

As well as interfaces for \textTeX{}, there are a small number of Lua functions provided here.

\begin{description}
\item[ltx.utils] Most public interfaces provided by the module are stored within the \texttt{ltx.utils} table.
\item[l3kernel] For compatibility reasons, there are also some deprecated interfaces provided in the \texttt{l3kernel} table. These do not return their result as Lua values but instead print them to \textTeX{}.
\item[l3kernel.charcat] \texttt{l3kernel.charcat(\langle charcode \rangle, \langle catcode \rangle)} Constructs a character of \langle charcode \rangle and \langle catcode \rangle and returns the result to \textTeX{}.
\item[l3kernel.elapsedtime] \texttt{l3kernel.elapsedtime()} Returns the CPU time in \langle scaled seconds \rangle since the start of the \textTeX{} 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.
\item[ltx.utils.filedump] \texttt{ltx.utils.filedump(\langle file \rangle, \langle offset \rangle, \langle length \rangle)} \texttt{l3kernel.filedump(\langle file \rangle, \langle offset \rangle, \langle length \rangle)} Returns the uppercase hexadecimal representation of the content of the \langle file \rangle read as bytes. If the \langle length \rangle is given, only this part of the file is returned; similarly, one may specify the \langle offset \rangle from the start of the file. If the \langle length \rangle is not given, the entire file is read starting at the \langle offset \rangle.
\item[ltx.utils.filemd5sum] \texttt{ltx.utils.filemd5sum(\langle file \rangle)} \texttt{l3kernel.filemd5sum(\langle file \rangle)} Returns 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 \textTeX{} behaviour. If the \langle file \rangle is not found, nothing is returned with \textit{no error raised}.
\end{description}
<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ltx.utils.filemoddate</code></td>
<td>Returns the date/time of last modification of the <code>&lt;file&gt;</code> in the format D: <code>&lt;year&gt;/&lt;month&gt;/&lt;day&gt;/&lt;hour&gt;/&lt;minute&gt;/&lt;second&gt;/&lt;offset&gt;</code> where the latter may be Z (UTC) or <code>&lt;plus-minus&gt;</code> <code>&lt;hours&gt;</code> ' <code>&lt;minutes&gt;</code>'. If the <code>&lt;file&gt;</code> is not found, nothing is returned with no error raised.</td>
</tr>
<tr>
<td><code>ltx.utils.filesize</code></td>
<td>Returns the size of the <code>&lt;file&gt;</code> in bytes. If the <code>&lt;file&gt;</code> is not found, nothing is returned with no error raised.</td>
</tr>
<tr>
<td><code>13kernel.filemoddate</code></td>
<td><code>&lt;file&gt;</code></td>
</tr>
<tr>
<td><code>13kernel.filesize</code></td>
<td><code>&lt;file&gt;</code></td>
</tr>
<tr>
<td><code>13kernel.resettimer</code></td>
<td>Resets the timer used by <code>13kernel.elapsetime</code>.</td>
</tr>
<tr>
<td><code>13kernel.shellescape</code></td>
<td>Executes the <code>&lt;cmd&gt;</code> and prints to the log as for pdfTeX.</td>
</tr>
<tr>
<td><code>13kernel.strcmp</code></td>
<td>Compares the two strings and returns 0 to \TeX{} if the two are identical.</td>
</tr>
</tbody>
</table>
Chapter 14

The \texttt{l3}legacy 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{} 2\epsilon 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 {⟨name⟩} {⟨true code⟩} {⟨false code⟩}
\end{verbatim}

Tests if the \LaTeX{} 2\epsilon/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}.

\begin{verbatim}
\legacy_if_set_true:n {⟨name⟩}
\legacy_if_set_false:n {⟨name⟩}
\end{verbatim}

Sets the \LaTeX{} 2\epsilon/plain \TeX{} conditional \texttt{\if⟨name⟩} (generated by \texttt{\newif}) to be \texttt{true} or \texttt{false}.

\begin{verbatim}
\legacy_if_set:nn {⟨name⟩} {⟨boolexpr⟩}
\end{verbatim}

Sets the \LaTeX{} 2\epsilon/plain \TeX{} conditional \texttt{\if⟨name⟩} (generated by \texttt{\newif}) to the result of evaluating the \langle boolean expression\rangle.
Part IV

Data types
Chapter 15

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:

\foo:n \{ a collection of \tokens \}

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

\foo:N \l\_some\_tl

In both cases, functions are available to test and manipulate the lists of tokens, and these have the module prefix \texttt{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 \texttt{\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 $\mathbb{N}$ argument, or $\mathbb{L}$, $\{$, or $\}$ (assuming normal \TeX{} category codes). Thus for example

\{ Hello \} ~ world

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

15.1 Creating and initialising token list variables

\tl_new:N \tl_new:N \tl\tl_var
\tl_new:c

Creates a new \texttt{(tl var)} or raises an error if the name is already taken. The declaration is global. The \texttt{(tl var)} is initially empty.
\tl_const:Nn  ⟨tl var⟩ \{<token list>\}

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⟩
\tl_clear:c
\tl_gclear:N
\tl_gclear:c

Clears all entries from the ⟨tl var⟩.

\tl_clear_new:N  ⟨tl var⟩
\tl_clear_new:c
\tl_gclear_new:N
\tl_gclear_new:c

Ensures that the ⟨tl var⟩ exists globally by applying \tl_new:N if necessary, then applies \tl_(g)clear:N to leave the ⟨tl var⟩ empty.

\tl_set_eq:NN  ⟨tl var⟩1 ⟨tl var⟩2
\tl_set_eq: (cN | Nc | cc)
\tl_gset_eq:NN  ⟨tl var⟩1 ⟨tl var⟩2
\tl_gset_eq: (cN | Nc | cc)

Sets the content of ⟨tl var⟩1 equal to that of ⟨tl var⟩2.

\tl_concat:NNN  ⟨tl var⟩1 ⟨tl var⟩2 ⟨tl var⟩3
\tl_concat:ccc
\tl_gconcat:NNN  ⟨tl var⟩1 ⟨tl var⟩2 ⟨tl var⟩3
\tl_gconcat:ccc

Concatenates the content of ⟨tl var⟩2 and ⟨tl var⟩3 together and saves the result in ⟨tl var⟩1. The ⟨tl var⟩2 is placed at the left side of the new token list.

\tl_if_exist_p:N  ⟨tl var⟩
\tl_if_exist:NTF  ⟨tl var⟩ \{<true code>\} \{<false code>\}
\tl_if_exist_p:c
\tl_if_exist:NTF
\tl_if_exist:c

Tests whether the ⟨tl var⟩ is currently defined. This does not check that the ⟨tl var⟩ really is a token list variable.

15.2 Adding data to token list variables

\tl_set:Nn
\tl_set:(NV|Nv|No|Nf|Nx|cn|cV|cV|cV|co|cf|cx)
\tl_gset:Nn
\tl_gset:(NV|Nv|No|Nf|Nx|cn|cV|cV|co|cf|cx)

Sets ⟨tl var⟩ to contain ⟨tokens⟩, removing any previous content from the variable.

\tl_put_left:Nn
\tl_put_left:(NV|Nv|No|Nf|Nx|cn|cV|co|cx)
\tl_gput_left:Nn
\tl_gput_left:(NV|Nv|No|Nf|Nx|cn|cV|co|cx)

Appends ⟨tokens⟩ to the left side of the current content of ⟨tl var⟩.

\tl_put_right:Nn
\tl_put_right:(NV|No|Nx|cn|cV|co|cx)
\tl_gput_right:Nn
\tl_gput_right:(NV|No|Nx|cn|cV|co|cx)

Appends \textit{tokens} to the right side of the current content of \textit{tl var}.

\section*{15.3 Modifying token list variables}

\tl_replace_once:Nnn \tl_replace_once:cn
\tl_greplace_once:Nnn \tl_greplace_once:cn

Updated: 2011-08-11

\tl_replace_all:Nnn \tl_replace_all:cn
\tl_greplace_all:Nnn \tl_greplace_all:cn

Updated: 2011-08-11

\tl_remove_once:Nn \tl_remove_once:cn
\tl_gremove_once:Nn \tl_gremove_once:cn

Updated: 2011-08-11

\tl_remove_all:Nn \tl_remove_all:cn
\tl_gremove_all:Nn \tl_gremove_all:cn

Updated: 2011-08-11

\tl_replace_once:Nnn \tl_replace_once:cn \tl_greplace_once:Nnn \tl_greplace_once:cn

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

\tl_replace_all:Nnn \tl_replace_all:cn \tl_greplace_all:Nnn \tl_greplace_all:cn

Replaces all occurrences of \textit{old tokens} in the \textit{tl var} with \textit{new tokens}. \textit{Old tokens} 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 \textit{old tokens} may remain after the replacement (see \tl_remove_all:Nn for an example).

\tl_remove_once:Nn \tl_remove_once:cn \tl_gremove_once:Nn \tl_gremove_once:cn

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

\tl_remove_all:Nn \tl_remove_all:cn \tl_gremove_all:Nn \tl_gremove_all:cn

Removes all occurrences of \textit{tokens} from the \textit{tl var}. \textit{Tokens} 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 \textit{tokens} may remain after the removal, for instance,

\tl_set:Nn \l_tmpa_tl {abbcdd} \tl_remove_all:Nn \l_tmpa_tl \{bc\}

results in \l_tmpa_tl containing abcd.

\section*{15.4 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}).
Sets \( \texttt{tl var} \) to contain \( \texttt{tokens} \), applying the category code régime specified in the \( \texttt{setup} \) before carrying out the assignment. (Category codes applied to tokens not explicitly covered by the \( \texttt{setup} \) are those in force at the point of use of \( \texttt{tl_set_rescan:Nnn} \).) This allows the \( \texttt{tl var} \) to contain material with category codes other than those that apply when \( \texttt{tokens} \) are absorbed. The \( \texttt{setup} \) is run within a group and may contain any valid input, although only changes in category codes, such as uses of \( \texttt{cctab_select:N} \), are relevant. See also \( \texttt{tl_set_rescan:Nnn} \).

\textbf{\texttt{TeX} hackers note:} The \( \texttt{tokens} \) 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 \( \texttt{setup} \)), 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.

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

\textbf{\texttt{TeX} hackers note:} The \( \texttt{tokens} \) 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 \( \texttt{setup} \)), 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.

Contrarily to the \( \texttt{scantokens} \) primitive, \( \texttt{tl_rescan:nn} \) tokenizes the whole string in the same category code regime rather than one token at a time, so that directives such as \( \texttt{verb} \) that rely on changing category codes will not function properly.
15.5 Token list conditionals

Tests if the \{token list\} consists only of blank spaces (i.e. contains no item). The test is true if \{token list\} is zero or more explicit space characters (explicit tokens with character code 32 and category code 10), and is false otherwise.

Tests if the \{token list\} 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 \l_tmpb_tl \{ abc \}
\tl_set:Nx \l_tmpb_tl { \tl_to_str:n \{ abc \} }
\tl_if_eq:nnTF \l_tmpa_tl \l_tmpb_tl \{ false \}
\end{verbatim}
yields false. See also \str_if_eq:nnTF for a comparison that ignores category codes.

Tests if the \{token list\} contains no item). The test is

Tests if \{token list\} and \{token list\} contain the same list of tokens, both in respect of character codes and category codes. This conditional is not expandable: see \tl_if_eq:NNTF for an expandable version when both token lists are stored in variables, or \str_if_eq:nNTF if category codes are not important.

Tests if \{token list\} and \{token list\} contain the same list of tokens, both in respect of character codes and category codes. This conditional is not expandable: see \tl_if_eq:NNTF for an expandable version when token lists are stored in variables, or \str_if_eq:nNTF if category codes are not important.

Tests if \{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 \( \langle \text{token list}_2 \rangle \) is found inside \( \langle \text{token list}_1 \rangle \). The \( \langle \text{token list}_2 \rangle \) 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 \( \langle \text{token list} \rangle \) 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 \( \langle \text{token list} \rangle \) 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 \tl\_count: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 \tl\_if\_eq: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 \tl\_case:Nn, which does nothing if there is no match, is also available.
15.6 Mapping over token lists

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

\texttt{\tl_map_function:NN} \langle \text{tl var} \rangle \langle \text{function} \rangle

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

\texttt{\tl_map_function:nN} \langle \text{token list} \rangle \langle \text{function} \rangle

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

\texttt{\tl_map_inline:Nn} \langle \text{tl var} \rangle \{ \langle inline function \rangle \}

Applies the \langle inline function \rangle to every \langle item \rangle stored within the \langle tl var \rangle. The \langle inline function \rangle should consist of code which receives the \langle item \rangle as #1. See also \texttt{\tl_map_\ldots function:nN}.

\texttt{\tl_map_inline:nn} \langle \text{token list} \rangle \{ \langle inline function \rangle \}

Applies the \langle inline function \rangle to every \langle item \rangle stored within the \langle token list \rangle. The \langle inline function \rangle should consist of code which receives the \langle item \rangle as #1. See also \texttt{\tl_map_\ldots function:nN}.

\texttt{\tl_map_tokens:Nn} \langle \text{tl var} \rangle \{ \langle code \rangle \}
\texttt{\tl_map_tokens:nn} \langle \text{tokens} \rangle \{ \langle code \rangle \}

Analogue of \texttt{\tl_map_function:NN} which maps several tokens instead of a single function. The \langle code \rangle receives each item in the \langle tl var \rangle or in \langle tokens \rangle as a trailing brace group. For instance,

\texttt{\tl_map_tokens:Nn} \l_my_tl \{ \prg_replicate:nn \{ 2 \} \}

expands to twice each item in the \langle tl var \rangle: for each item in \l_my_tl the function \texttt{\prg_replicate:nn} receives 2 and \langle item \rangle as its two arguments. The function \texttt{\tl_map_inline:Nn} is typically faster but is not expandable.

\texttt{\tl_map_variable:NNn} \langle \text{tl var} \rangle \langle \text{variable} \rangle \{ \langle code \rangle \}

Stores each \langle item \rangle of the \langle tl var \rangle in turn in the \langle token list \rangle \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 item \rangle in the \langle tl var \rangle, or its original value if the \langle tl var \rangle is blank. See also \texttt{\tl_map_inline:Nn}. 
\tl_map_variable:nNn \tl_map_variable:nNn \{(token list}\} \{variable\} \{(code)\}
Stores each \langle item \rangle of the \langle token list \rangle in turn in the \langle token list \rangle \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 item \rangle in the \langle tl var \rangle, or its original value if the \langle tl var \rangle is blank. See also \tl_map_inline:nNn.

\tl_map_break: \tl_map_break:
Updated: 2012-06-29
Used to terminate a \tl_map_... function before all entries in the \langle token list variable \rangle have been processed. This normally takes place within a conditional statement, for example

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

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

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

\tl_map_break:n \tl_map_break:n \langle code \rangle
Updated: 2012-06-29
Used to terminate a \tl_map_... function before all entries in the \langle token list variable \rangle
have been processed, inserting the \langle code \rangle after the mapping has ended. This normally
takes place within a conditional statement, for example

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

Use outside of a \tl_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.
15.7 Using token lists

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

\textbf{\textsc{\TeX}hacker note:} This is the \texttt{\textsc{\TeX}} primitive \texttt{\detokenize}. Converting a \langle\text{token list}\rangle to a \langle\text{string}\rangle yields a concatenation of the string representations of every token in the \langle\text{token list}\rangle. The string representation of a control sequence is

- an escape character, whose character code is given by the internal parameter \texttt{\escapechar}, absent if the \texttt{\escapechar} is negative or greater than the largest character code;
- the control sequence name, as defined by \texttt{\cs_to_str:N};
- a space, unless the control sequence name is a single character whose category at the time of expansion of \texttt{\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 \texttt{\escapechar}: for instance \texttt{\tl_to_str:n \{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 \texttt{a} is currently not a letter.

\tl_to_str:N \to \tl_to_str:N \tl var

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

\tl_use:N \to \tl_use:N \tl var

Recover the content of a \langle\text{tl 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\text{tl var}\rangle directly without an accessor function.

15.8 Working with the content of token lists

\tl_count:n \to \tl_count:n \{\langle\text{tokens}\rangle\}

Counts the number of \langle\text{items}\rangle in \langle\text{tokens}\rangle 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 \langle\text{tokens}\rangle. See also \texttt{\tl_count:N}. This function requires three expansions, giving an \langle\text{integer denotation}\rangle.
\tl_count:N \tl var

Counts the number of token groups in the \langle tl var \rangle 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 \langle tl var \rangle. See also \tl_count:n. This function requires three expansions, giving an \langle integer denotation \rangle.

\tl_count_tokens:n \langle tokens \rangle

Counts the number of TEX tokens in the \langle tokens \rangle 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 \langle \langle token list \rangle \rangle

Reverses the order of the \langle items \rangle in the \langle token list \rangle, so that \langle item_1 \rangle\langle item_2 \rangle\langle item_3 \rangle \ldots \langle item_n \rangle becomes \langle item_n \rangle \ldots \langle item_3 \rangle\langle item_2 \rangle\langle item_1 \rangle. This process preserves unprotected space within the \langle token list \rangle. 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.

**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_reverse_items:n \langle \langle token list \rangle \rangle

Reverses the order of the \langle items \rangle stored in \langle tl var \rangle, so that \langle item_1 \rangle\langle item_2 \rangle\langle item_3 \rangle \ldots \langle item_n \rangle becomes \langle item_n \rangle \ldots \langle item_3 \rangle\langle item_2 \rangle\langle item_1 \rangle. This process preserves unprotected spaces within the \langle token list variable \rangle. Braced token groups are copied without reversing the order of tokens, but keep the outer set of braces. See also \tl_reverse:n, and, for improved performance, \tl_reverse_items:n.

**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:n \langle \langle token list \rangle \rangle

Removes any leading and trailing explicit space characters (explicit tokens with character code 32 and category code 10) from the \langle token list \rangle 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 \tl trim_spaces_apply:oN

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

\tl trim_spaces:N \tl trim_spaces:c \tl gtrim_spaces:N \tl gtrim_spaces:c

Removes any leading and trailing explicit space characters (explicit tokens with character code 32 and category code 10) from the content of the \textit{tl var}. Note that this therefore \textit{resets} the content of the variable.

\tl sort:Nn \tl sort:cn \tl gsort:Nn \tl gsort:cn

Sorts the items in the \textit{tl var} according to the \textit{comparison code}, and assigns the result to \textit{tl var}. The details of sorting comparison are described in Section 6.1.

\tl sort:nN \tl sort:n \tl sort:cn

Sorts the items in the \textit{token list}, using the \textit{conditional} to compare items, and leaves the result in the input stream. The \textit{conditional} should have signature :nnTF, and return \textbf{true} if the two items being compared should be left in the same order, and \textbf{false} if the items should be swapped. The details of sorting comparison are described in Section 6.1.

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

15.9 \textbf{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.
\texttt{\tl_head:n \{} \langle \text{token list} \rangle \}

Leaves in the input stream the first \langle item \rangle in the \langle token list \rangle, discarding the rest of the \langle token list \rangle. All leading explicit space characters (explicit tokens with character code 32 and category code 10) are discarded; for example

\texttt{\tl_head:n \{ abc \}}

and

\texttt{\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

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

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

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

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

Leaves in the input stream the first \langle item \rangle in the \langle token list \rangle, discarding the rest of the \langle token list \rangle. All leading explicit space characters (explicit tokens with character code 32 and category code 10) are discarded. A blank \langle token list \rangle (which consists only of space characters) results in a low-level \LaTeX 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, \texttt{\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, \texttt{\tl_head:n} should be preferred if the number of expansions is not critical.

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

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

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

and

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

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

\textbf{LaTeX hackers note:} The result is returned within \texttt{\exp_not:n}, which means that the token list does not expand further when appearing in an x-type argument expansion.
\tl_if_head_eq_catcode_p:nN \tl_if_head_eq_catcode_p:oN \tl_if_head_eq_catcode:nNTF
⟨token list⟩ ⟨test token⟩ {⟨true code⟩} {⟨false code⟩}

Tests if the first (token) in the (token list) has the same category code as the (test token).
In the case where the (token list) is empty, the test is always false.

\tl_if_head_eq_charcode_p:nN \tl_if_head_eq_charcode_p:fN \tl_if_head_eq_charcode:nN \tl_if_head_eq_charcode:fN
⟨token list⟩ ⟨test token⟩ {⟨true code⟩} {⟨false code⟩}

Tests if the first (token) in the (token list) has the same character code as the (test token).
In the case where the (token list) is empty, the test is always false.

\tl_if_head_eq_meaning_p:nN \tl_if_head_eq_meaning:nN \tl_if_head_eq_meaning_p:oN \tl_if_head_eq_meaning:oN
⟨token list⟩ ⟨test token⟩ {⟨true code⟩} {⟨false code⟩}

Tests if the first (token) in the (token list) has the same meaning as the (test token). In the case where (token list) is empty, the test is always false.

\tl_if_head_is_group_p:n \tl_if_head_is_group:nTF
\tl_if_head_is_N_type_p:n \tl_if_head_is_N_type:nTF
\tl_if_head_is_space_p:n \tl_if_head_is_space:nTF

Tests if the first (token) in the (token list) is an explicit begin-group character (with category code 1 and any character code), in other words, if the (token list) starts with a brace group. In particular, the test is false if the (token list) starts with an implicit token such as \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 (token) in the (token list) is a normal 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 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 (token) in the (token list) is an explicit space character (explicit token with character code 12 and category code 10). In particular, the test is false if the (token list) starts with an implicit token such as \c_space_token, or if it is empty. This function is useful to implement actions on token lists on a token by token basis.
15.10 Using a single item

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

Indexing items in the \textit{token list} from 1 on the left, this function evaluates the \textit{integer expression} and leaves the appropriate item from the \textit{token list} in the input stream. If the \textit{integer expression} 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{\TeX}hackers note: The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp_not:n}), which means that the \textit{item} does not expand further when appearing in an \texttt{x}-type argument expansion.

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

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

\textbf{\TeX}hackers note: The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp_not:n}), which means that the \textit{item} does not expand further when appearing in an \texttt{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 \textit{actual start point} is determined as \( M = m \) if \( m > 0 \) and as \( M = l + m + 1 \) if \( m < 0 \). Similarly the \textit{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

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

Here are some more interesting examples. The calls

\begin{verbatim}
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { 2 } { 5 } }
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { 2 } { -3 } }
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { -6 } { 5 } }
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { -6 } { -3 } }
\end{verbatim}

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

\begin{verbatim}
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { 2 } { 5 } }
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { 2 } { -3 } }
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { -6 } { 5 } }
\iow_term:x { \tl_range:nnn { abcd-{e{}}fg } { -6 } { -3 } }
\end{verbatim}

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

\begin{verbatim}
\tl_range:nnn { abcd-{e{}}f } { 2 } { 4 }
\end{verbatim}

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 { \langle start \rangle } { 3 } { -1 }}. Similarly, for discarding the last item, we can do \texttt{\tl_range:nnn { \langle start \rangle } { 1 } { -2 }}.

For better performance, see \texttt{\tl_range_braced:nnn} and \texttt{\tl_range_unbraced:nnn}.

\textbf{\TeX}Hackers 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.
15.11 Viewing token lists

\tl_show:N  
\tl_show:c  
\tl_log:N  
\tl_log:c  
\tl_log:n  

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

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

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

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

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

\tl_log:n \{\langle token list \rangle\}  
Writes the \langle token list \rangle in the log file. See also \tl_show:n which displays the result in the terminal.  

15.12 Constant token lists

\c_empty_tl  
Constant that is always empty.  

\c_novalue_tl  
New: 2017-11-14  
A marker for the absence of an argument. This constant tl can safely be typeset (cf. \texttt{q-nil}), with the result being \texttt{-NoValue-}. It is important to note that \c_novalue_tl is constructed such that it will not match the simple text input \texttt{-NoValue-}, \textit{i.e.} that  

\tl_if_eq:NnTF \c_novalue_tl \{ -NoValue- \}  
is logically \texttt{false}. The \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 tl.  

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

\l_tmpa_tl \l_tmpb_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.

\g_tmpa_tl \g_tmpb_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.
Chapter 16

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{\textbackslash str\ldots:N}, which expect a token list or string variable as their argument;
- \texttt{\textbackslash str\ldots:n}, taking any token list (or string) as an argument;
- \texttt{\textbackslash str\ldots: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.
16.1 Building strings

\textbf{\texttt{\string_new:N}}
\texttt{\string_new:c}
\texttt{\string_clear:N}
\texttt{\string_clear:c}
\texttt{\string_gclear:N}
\texttt{\string_gclear:c}
\texttt{\string_set_eq:NN}
\texttt{\string_set_eq:cc}
\texttt{\string_gset_eq:NN}
\texttt{\string_gset_eq:cc}
\texttt{\string_concat:NNN}
\texttt{\string_concat:ccc}
\texttt{\string_gconcat:NNN}
\texttt{\string_gconcat:ccc}

\textbf{\texttt{\string_new:N}} \langle \text{str var} \rangle
\textbf{\texttt{\string_new:c}}

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.

\textbf{\texttt{\string_const:Nn}}
\texttt{\string_const:Nn}
\texttt{\string_const:cc}
\texttt{\string_gset_eq:NN}
\texttt{\string_gset_eq:cc}
\texttt{\string_concat:NNN}
\texttt{\string_concat:ccc}
\texttt{\string_gconcat:NNN}
\texttt{\string_gconcat:ccc}

\textbf{\texttt{\str_const:Nn}} \langle \text{str var} \rangle \{ \langle \text{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 \text{token list} \rangle, converted to a string.

\textbf{\texttt{\str_clear:N}}
\texttt{\str_clear:c}
\texttt{\str_gclear:N}
\texttt{\str_gclear:c}
\texttt{\str_set_eq:NN}
\texttt{\str_set_eq:cc}
\texttt{\str_gset_eq:NN}
\texttt{\str_gset_eq:cc}
\texttt{\str_concat:NNN}
\texttt{\str_concat:ccc}
\texttt{\str_gconcat:NNN}
\texttt{\str_gconcat:ccc}

\textbf{\texttt{\str_new:N}} \langle \text{str var} \rangle
\textbf{\texttt{\str_new:c}}

Clears the content of the \langle str var \rangle.

\textbf{\texttt{\str_set_eq:NN}} \langle \text{str var} \rangle \langle \text{str var} \rangle
\textbf{\texttt{\str_set_eq:cc}} \langle \text{str var} \rangle \langle \text{str var} \rangle
\textbf{\texttt{\str_gset_eq:NN}} \langle \text{str var} \rangle \langle \text{str var} \rangle
\textbf{\texttt{\str_gset_eq:cc}} \langle \text{str var} \rangle \langle \text{str var} \rangle
\textbf{\texttt{\str_concat:NNN}} \langle \text{str var} \rangle \langle \text{str var} \rangle \langle \text{str var} \rangle
\textbf{\texttt{\str_concat:ccc}} \langle \text{str var} \rangle \langle \text{str var} \rangle \langle \text{str var} \rangle

\textbf{\texttt{\str_set:Nn}}
\texttt{\str_set:Nn}
\texttt{\str_set:Nn}
\texttt{\str_set:cc}
\texttt{\str_set:cc}
\texttt{\str_gset:Nn}
\texttt{\str_gset:Nn}
\texttt{\str_gset:cc}
\texttt{\str_gset:cc}

\textbf{\texttt{\str_set:Nn}} \langle \text{str var} \rangle \{ \langle \text{token list} \rangle \}

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

16.2 Adding data to string variables

\textbf{\texttt{\str_set:Nn}}
\texttt{\str_set:Nn}
\texttt{\str_set:Nn}
\texttt{\str_set:cc}
\texttt{\str_set:cc}
\texttt{\str_gset:Nn}
\texttt{\str_gset:Nn}
\texttt{\str_gset:cc}
\texttt{\str_gset:cc}

\textbf{\texttt{\str_set:Nn}} \langle \text{str var} \rangle \{ \langle \text{token list} \rangle \}

Converts the \langle \text{token list} \rangle to a \langle \text{string} \rangle, and stores the result in \langle str var \rangle.
16.3 Modifying string variables

\texttt{str_replace_once:Nnn} \str_replace_once:nnn \str_replace_once:cn \str_greplace_once:Nn \str_greplace_once:cn

\begin{verbatim}
\str_replace_once:Nnn \str_replace_once:nnn \str_replace_once:cn \str_greplace_once:Nnn \str_greplace_once:cn
\end{verbatim}

\begin{itemize}
  \item \texttt{str_replace_once:Nnn} \str_replace_once:nnn \str_replace_once:cn \str_greplace_once:Nnn \str_greplace_once:cn
  \item \texttt{str_replace_all:Nnn} \str_replace_all:nnn \str_replace_all:cn \str_greplace_all:Nnn \str_greplace_all:cn
  \item \texttt{str_remove_once:Nn} \str_remove_once:nn \str_gremove_once:Nn \str_gremove_once:cn
  \item \texttt{str_remove_all:Nn} \str_remove_all:nn \str_gremove_all:Nn \str_gremove_all:cn
\end{itemize}

\begin{itemize}
  \item \texttt{str_replace_all:Nnn} \str_replace_all:nnn \str_replace_all:cn \str_greplace_all:Nnn \str_greplace_all:cn
  \item \texttt{str_remove_all:Nn} \str_remove_all:nn \str_gremove_all:Nn \str_gremove_all:cn
\end{itemize}

\texttt{str_replace_once:Nnn} \str_replace_once:nnn \str_replace_once:cn \str_greplace_once:Nnn \str_greplace_once:cn

\begin{itemize}
  \item \texttt{str_replace_once:Nnn} \str_replace_once:nnn \str_replace_once:cn \str_greplace_once:Nnn \str_greplace_once:cn
  \item \texttt{str_replace_all:Nnn} \str_replace_all:nnn \str_replace_all:cn \str_greplace_all:Nnn \str_greplace_all:cn
  \item \texttt{str_remove_once:Nn} \str_remove_once:nn \str_gremove_once:Nn \str_gremove_once:cn
  \item \texttt{str_remove_all:Nn} \str_remove_all:nn \str_gremove_all:Nn \str_gremove_all:cn
\end{itemize}

\begin{itemize}
  \item \texttt{str_replace_once:Nnn} \str_replace_once:nnn \str_replace_once:cn \str_greplace_once:Nnn \str_greplace_once:cn
  \item \texttt{str_replace_all:Nnn} \str_replace_all:nnn \str_replace_all:cn \str_greplace_all:Nnn \str_greplace_all:cn
  \item \texttt{str_remove_once:Nn} \str_remove_once:nn \str_gremove_once:Nn \str_gremove_once:cn
  \item \texttt{str_remove_all:Nn} \str_remove_all:nn \str_gremove_all:Nn \str_gremove_all:cn
\end{itemize}

\begin{itemize}
  \item \texttt{str_replace_all:Nnn} \str_replace_all:nnn \str_replace_all:cn \str_greplace_all:Nnn \str_greplace_all:cn
  \item \texttt{str_remove_all:Nn} \str_remove_all:nn \str_gremove_all:Nn \str_gremove_all:cn
\end{itemize}

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

Converts the \langle \texttt{token list} \rangle to a \langle \texttt{string} \rangle, and appends the result to \langle \texttt{str var} \rangle. The current contents of the \langle \texttt{str var} \rangle are not automatically converted to a string.

\begin{verbatim}
\str_set:Nn \l_tmpa_str {abbccd} \str_remove_all:Nn \l_tmpa_str \l_tmpa_str {bc}
\end{verbatim}

results in \l_tmpa_str containing abcd.
## 16.4 String conditionals

\[\text{\texttt{\textbackslash str\_if\_exist\_p:}\texttt{\textbackslash N} \ast \texttt{\textbackslash str\_if\_exist\_p:}\texttt{\textbackslash C} \ast \texttt{\textbackslash str\_if\_exist:}\texttt{\textbackslash NTF} \ast \texttt{\textbackslash str\_if\_exist:}\texttt{\textbackslash CTF} \ast}\]

New: 2015-09-18

Tests whether the \texttt{(str var)} is currently defined. This does not check that the \texttt{(str var)} really is a string.

\[\text{\texttt{\textbackslash str\_if\_empty\_p:}\texttt{\textbackslash N} \ast \texttt{\textbackslash str\_if\_empty\_p:}\texttt{\textbackslash C} \ast \texttt{\textbackslash str\_if\_empty:}\texttt{\textbackslash NTF} \ast \texttt{\textbackslash str\_if\_empty:}\texttt{\textbackslash CTF} \ast}\]

New: 2015-09-18

Tests if the \texttt{(string variable)} is entirely empty \textit{(i.e. contains no characters at all)}.

\[\text{\texttt{\textbackslash str\_if\_eq\_p:}\texttt{\textbackslash NN} \ast \texttt{\textbackslash str\_if\_eq\_p:}\texttt{(Nc|cN|cc)} \ast \texttt{\textbackslash str\_if\_eq:}\texttt{\textbackslash NNTF} \ast \texttt{\textbackslash str\_if\_eq:}\texttt{(Nc|cN|cc)} \ast}\]

Updated: 2018-06-18

Compares the content of two \texttt{(str variables)} and is logically \texttt{true} if the two contain the same characters in the same order. See \texttt{\textbackslash tl\_if\_eq:nnTF} to compare tokens (including their category codes) rather than characters.

\[\text{\texttt{\textbackslash str\_if\_eq\_p:}\texttt{\textbackslash nn} \ast \texttt{\textbackslash str\_if\_eq\_p:}\texttt{(Vn|on|no|nV|vn|nv|ee)} \ast \texttt{\textbackslash str\_if\_eq:}\texttt{\textbackslash NN} \ast \texttt{\textbackslash str\_if\_eq:}\texttt{(Vn|on|no|nV|vn|nv|ee)} \ast}\]

Updated: 2018-06-18

Compares the two \texttt{(token lists)} on a character by character basis (namely after converting them to strings), and is \texttt{true} if the two \texttt{(strings)} contain the same characters in the same order. Thus for example

\[\texttt{\textbackslash str\_if\_eq\_p:}\texttt{\textbackslash no} \{\texttt{abc}\} \{\texttt{\textbackslash tl\_to\_str:n} \{\texttt{abc}\}\}\]

is logically \texttt{true}. See \texttt{\textbackslash tl\_if\_eq:nnTF} to compare tokens (including their category codes) rather than characters.

\[\text{\texttt{\textbackslash str\_if\_in:}\texttt{\textbackslash NnTF} \ast \texttt{\textbackslash str\_if\_in:}\texttt{\textbackslash NTF} \ast \texttt{\textbackslash str\_if\_in:}\texttt{\textbackslash cTF} \ast} \]

New: 2017-10-08

Converts the \texttt{(token list)} to a \texttt{(string)} and tests if that \texttt{(string)} is found in the content of the \texttt{(str var)}.

\[\text{\texttt{\textbackslash str\_if\_in:}\texttt{\textbackslash nnTF} \ast \texttt{\textbackslash str\_if\_in:}\texttt{\textbackslash nTF} \ast} \]

New: 2017-10-08

Converts both \texttt{(token lists)} to \texttt{(strings)} and tests whether \texttt{(string2)} is found inside \texttt{(string1)}. 

121
Compares the \textit{test string} in turn with each of the \textit{string cases} (all token lists are converted to strings). If the two are equal (as described for \texttt{\textbackslash str\_if\_eq:nnTF}) 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{\textbackslash str\_case:nn}, which does nothing if there is no match, is also available.

\bigskip

\texttt{\textbackslash str\_case:nn} \textdagger
\texttt{\textbackslash str\_case:(Vn|on|nV|nv)} \textdagger
\texttt{\textbackslash str\_case:nnTF} \textdagger
\texttt{\textbackslash str\_case:(Vn|on|nV|nv)TF} \textdagger

\texttt{\textbackslash str\_case:e:nn} \textdagger
\texttt{\textbackslash str\_case:e:nnTF} \textdagger

\texttt{\textbackslash str\_case:nn} \textdagger
\texttt{\textbackslash str\_case:(Vn|on|nV|nv)} \textdagger
\texttt{\textbackslash str\_case:nnTF} \textdagger
\texttt{\textbackslash str\_case:(Vn|on|nV|nv)TF} \textdagger

\texttt{\textbackslash str\_case_e:nn} \textdagger
\texttt{\textbackslash str\_case_e:nnTF} \textdagger

\texttt{\textbackslash str\_case_e:nn} \textdagger
\texttt{\textbackslash str\_case_e:nnTF} \textdagger

New: 2013-07-24
Updated: 2015-02-28

New: 2018-06-19

Compares the \textit{test string} in turn with each of the \textit{string cases} (all token lists are converted to strings). If the two are equal (as described for \texttt{\textbackslash str\_if\_eq:nnTF}) 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{\textbackslash str\_case:nn}, which does nothing if there is no match, is also available. The \textit{test string} is expanded in each comparison, and must always yield the same result: for example, random numbers must not be used within this string.
\str_compare_p:nNn \str_compare_p:eNe \str_compare:nNnTF \str_compare:eNeTF

Revised: 2021-05-17

\str_compare_p:nNn {\{tl_1\}} \langle relation \rangle \{\{tl_2\}\}
\str_compare:pNnTF {\{tl_1\}} \langle relation \rangle \{\{tl_2\}\} \{\langle true code \rangle \} \{\langle false code \rangle \}

Compares the two \langle token lists \rangle on a character by character basis (namely after converting them to strings) in a lexicographic order according to the character codes of the characters. The \langle relation \rangle can be \langle < \rangle, \langle = \rangle, or \langle > \rangle and the test is \langle true \rangle under the following conditions:

- for \langle < \rangle, if the first string is earlier than the second in lexicographic order;
- for \langle = \rangle, if the two strings have exactly the same characters;
- for \langle > \rangle, if the first string is later than the second in lexicographic order.

Thus for example the following is logically \langle true \rangle:
\str_compare_p:nNn { ab } < \{ abc \}

\TeX{} hackers note: This is a wrapper around the \TeX{} primitive \langle (pdf)strcmp \rangle. It is meant for programming and not for sorting textual contents, as it simply considers character codes and not more elaborate considerations of grapheme clusters, locale, etc.

16.5 Mapping over strings

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.

\str_map_function:nN \str_map_function:NN \str_map_function:cN

Revised: 2017-11-14

\str_map_function:nN \{\langle token list \rangle\} \langle function \rangle
\str_map_function:NN \{\langle str var \rangle\} \{\langle function \rangle \}

Converts the \langle token list \rangle to a \langle string \rangle then applies \langle function \rangle to every \langle character \rangle in the \langle string \rangle including spaces.

\str_map_function:nN \str_map_function:NN \str_map_function:cn

\str_map_inline:nn \str_map_inline:Nn \str_map_inline:cn

Revised: 2017-11-14

\str_map_inline:nn \{\langle token list \rangle\} \{\langle inline function \rangle \}
\str_map_inline:Nn \{\langle str var \rangle\} \{\langle inline function \rangle \}

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 str var \rangle including spaces. The \langle inline function \rangle should consist of code which receives the \langle character \rangle as \#1.

\str_map_variable:nNn \str_map_variable:NNn \str_map_variable:cNn

Revised: 2017-11-14

\str_map_variable:nNn \{\langle token list \rangle\} \{variable\} \{\langle code \rangle\}
\str_map_variable:NNn \{\langle str var \rangle\} \{variable\} \{\langle code \rangle\}

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 \langle string or token list \rangle \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.
Converting the \token list to a \string then applies \code to every \character in the \string including spaces. The \code receives each character as a trailing brace group. This is equivalent to \str_map_function:nN if the \code consists of a single function.

\str_map_break:

Used to terminate a \str_map_... function before all characters in the \string have been processed. 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: }\
  % Do something useful\
}

See also \str_map_break:n. 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 continuing with the code that follows the loop. This depends on the design of the mapping function.

\str_map_break:n

Used to terminate a \str_map_... function before all characters in the \string have been processed, inserting the \code 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 \code is inserted into the input stream. This depends on the design of the mapping function.

\str_use:N

Recovers the content of a \str 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 \str directly without an accessor function.

16.6 Working with the content of strings
Leaves in the input stream the number of characters in the string representation of ⟨token list⟩, as an integer denotation. The functions differ in their treatment of spaces. In the case of \texttt{str\_count:N} and \texttt{str\_count:n}, all characters including spaces are counted. The \texttt{str\_count\_ignore\_spaces:n} function leaves the number of non-space characters in the input stream.

Leaves in the input stream the number of space characters in the string representation of ⟨token list⟩, as an integer denotation. Of course, this function has no \_ignore\_spaces variant.

Converts the ⟨token list⟩ into a ⟨string⟩. The first character in the ⟨string⟩ is then left in the input stream, with category code “other”. The functions differ if the first character is a space: \texttt{str\_head:N} and \texttt{str\_head:n} return a space token with category code 10 (blank space), while the \texttt{str\_head\_ignore\_spaces:n} function ignores this space character and leaves the first non-space character in the input stream. If the ⟨string⟩ is empty (or only contains spaces in the case of the \_ignore\_spaces function), then nothing is left on the input stream.

Converts the ⟨token list⟩ to a ⟨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{str\_tail:N} and \texttt{str\_tail:n} only trim that space, while \texttt{str\_tail\_ignore\_spaces:n} removes the first non-space character and any space before it. If the ⟨token list⟩ is empty (or blank in the case of the \_ignore\_spaces variant), then nothing is left on the input stream.
\str_item:Nn \str_item:nn \{\token\ list\} \{\integer\ expression\}
\str_item:nn
\str_item_ignore_spaces:nn

New: 2015-09-18

Converts the \{\token\ list\} to a \{\string\}, and leaves in the input stream the character in position \{\integer\ expression\} of the \{\string\}, starting at 1 for the first (left-most) character. In the case of \str_item:Nn and \str_item:nn, all characters including spaces are taken into account. The \str_item_ignore_spaces:nn function skips spaces when counting characters. If the \{\integer\ expression\} is negative, characters are counted from the end of the \{\string\}. Hence, \(-1\) is the right-most character, etc.

\str_range:nnn \str_range:nn \str_range:cnn \str_range:nnn \str_range_ignore_spaces:nnn

New: 2015-09-18

Converts the \{\token\ list\} to a \{\string\}, and leaves in the input stream the characters from the \{\start\ index\} to the \{\end\ index\} inclusive. Spaces are preserved and counted as items (contrast this with \tl_range:nnn where spaces are not counted as items and are possibly discarded from the output).

Here \{\start\ index\} and \{\end\ index\} 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,

\io_term:x { \str_range:nnn \{abcdef\} \{2\} \{5\} }
\io_term:x { \str_range:nnn \{abcdef\} \{-4\} \{-1\} }
\io_term:x { \str_range:nnn \{abcdef\} \{-2\} \{-1\} }
\io_term:x { \str_range:nnn \{abcdef\} \{0\} \{-1\} }

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

\io_term:x { \str_range:nnn \{abcdef\} \{5\} \{2\} }
\io_term:x { \str_range:nnn \{abcdef\} \{-1\} \{-4\} }

both yield empty strings.

The behavior of \str_range_ignore_spaces:nnn is similar, but spaces are removed before starting the job. The input

\io_term:x { \str_range:nnn \{abcdef\} \{2\} \{5\} }
\io_term:x { \str_range:nnn \{abcdef\} \{2\} \{-3\} }

126
will print four instances of `bcde`, four instances of `bc e` and eight instances of `bcde`. 
16.7 String manipulation

\str_lowercase:n \str_lowercase:f \str_uppercase:n \str_uppercase:f

Converts the input \{\text{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
\str_uppercase:f \str_lowercase:f \tl_head:n \tl_tail:n
\myfunc:nn \cs_new_protected:Npn \myfunc:nn #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, \text_uppercase:n and \text_titlecase:n functions which correctly deal with context-dependence and other factors appropriate to text case changing.

\TeX\text{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 Xe\TeX\ and Lua\TeX.
\texttt{\textbackslash{}str\_foldcase:n} \{\langle \text{tokens} \rangle \}

Converts the input \langle \text{tokens} \rangle to their string representation, as described for \texttt{\textbackslash{}tl\_to\_str:n}, and then folds the case of the resulting \langle \text{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 \texttt{\textbackslash{}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 \texttt{\textbackslash{}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. İ always folds to i and not to ĵ).

\textbf{T\TeX{}hackers note}: As with all \texttt{expl3} functions, the input supported by \texttt{\textbackslash{}str\_foldcase:n} is \texttt{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 \texttt{\textbackslash{}tl\_to\_str:n}.

### 16.8 Viewing strings

\texttt{\textbackslash{}str\_show:N} \langle \text{str var} \rangle

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

\texttt{\textbackslash{}str\_log:N} \langle \text{str var} \rangle

Writes the content of the \langle \text{str var} \rangle in the log file.
16.9 Constant token lists

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

\c_ampersand_str
\c_atsign_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
\c_zero_str

16.10 Scratch strings

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.

\l_tmpa_str
\l_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.

\g_tmpa_str
\g_tmpb_str
Chapter 17

The l3str-convert package: string encoding conversions

17.1 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.\(^6\)

- Bytes are translated to \(\LaTeX\) 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.\(^6\)

---

\(^6\)Encodings and escapings will be added as they are requested.
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, latin1</td>
<td>ISO 8859-1</td>
</tr>
<tr>
<td>iso88592, latin2</td>
<td>ISO 8859-2</td>
</tr>
<tr>
<td>iso88593, latin3</td>
<td>ISO 8859-3</td>
</tr>
<tr>
<td>iso88594, latin4</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, latin5</td>
<td>ISO 8859-9</td>
</tr>
<tr>
<td>iso885910, latin6</td>
<td>ISO 8859-10</td>
</tr>
<tr>
<td>iso885911</td>
<td>ISO 8859-11</td>
</tr>
<tr>
<td>iso885913, latin7</td>
<td>ISO 8859-13</td>
</tr>
<tr>
<td>iso885914, latin8</td>
<td>ISO 8859-14</td>
</tr>
<tr>
<td>iso885915, latin9</td>
<td>ISO 8859-15</td>
</tr>
<tr>
<td>iso885916, latin10</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>
<tr>
<td>default</td>
<td>Like utf8 with 8-bit engines, and like native with unicode-engines</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 \pdfescapearname</td>
</tr>
<tr>
<td>string</td>
<td>See \pdfescapestring</td>
</tr>
<tr>
<td>url</td>
<td>Encoding used in URLs</td>
</tr>
</tbody>
</table>
17.2 Conversion functions

\str_set_convert:Nnnn \str_gset_convert:Nnnn

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

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,

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

results in the variable \l_foo_str holding the string FEFF00480065006C006F0021.

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 ⟨encoding 1⟩ and ⟨encoding 2⟩, 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.

\str_set_convert:NnnnTF \str_gset_convert:NnnnTF

\str_set_convert:NnnnTF (str var) {⟨string⟩} {⟨name 1⟩} {⟨name 2⟩} {⟨true code⟩} {⟨false code⟩}

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

17.3 Conversion by expansion (for PDF contexts)

A small number of expandable functions are provided for use in PDF string/name contexts. These assume UTF-8 and no escaping in the input.

\str_convert_pdfname:n *

\str_convert_pdfname:n (⟨string⟩)

As \str_convert_pdfname:n, converts the ⟨string⟩ on a byte-by-byte basis with non-ASCII codepoints escaped using hashes.

17.4 Possibilities, and things to do

Encoding/escaping-related tasks.
• In Xe\TeX/Lua\TeX, would it be better to use the \ldots 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 \texttt{scantokens}.

• Change \texttt{\texttt{\texttt{\texttt{str_set_convert:Nnnn}}} to expand its last two arguments.

• Describe the internal format in the code comments. Refuse code points in [\texttt{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 \texttt{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 \texttt{stringenc}). CESU?

• More escapings: \texttt{ascii85}, shell escapes, lua escapes, etc.?
Chapter 18

The \texttt{I3quark} 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}.

18.1 Quarks

Quarks are control sequences (and in fact, token lists) 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.
18.2 Defining quarks

\quark_new:N \quark_new:N \quark

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

\q_stop

Used as a marker for delimited arguments, such as

\cs_set:Npn \tmp:w #1#2 \q_stop \{#1\}

\q_mark

Used as a marker for delimited arguments when \q_stop is already in use.

\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 \q_stop, which is only ever used as a delimiter).

\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 \prop_get:NnN if there is no data to return.

18.3 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.

\quark_if_nil_p:N * \quark_if_nil_p:N \token \quark_if_nil_p:n \token \{\true code\} \{\false code\}

Tests if the \token is equal to \q_nil.

\quark_if_nil:nTF * \quark_if_nil:nTF \token \{\true code\} \{\false code\}

Tests if the \token list contains only \q_nil (distinct from \token list being empty or containing \q_nil plus one or more other tokens).

\quark_if_no_value_p:N * \quark_if_no_value_p:n \token \quark_if_no_value:p:c \{\true code\} \{\false code\}

Tests if the \token is equal to \q_no_value.

\quark_if_no_value:nTF * \quark_if_no_value:nTF \token \{\true code\} \{\false code\}

Tests if the \token list contains only \q_no_value (distinct from \token list being empty or containing \q_no_value plus one or more other tokens).
18.4 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 18.4.1.

\_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.

\_q\_recursion\_stop

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 token \rangle

Tests if \langle 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 \* \quark\_if\_recursion\_tail\_stop:n \{\langle token \rangle\}

\quark\_if\_recursion\_tail\_stop:o \* \quark\_if\_recursion\_tail\_stop:o

Tests if \langle token list \rangle contains only \_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 insertion \rangle code is then added to the input stream after the recursion has ended.

\quark\_if\_recursion\_tail\_stop:nn \* \quark\_if\_recursion\_tail\_stop:nn \langle\langle token list\rangle\rangle \langle\langle insertion\rangle\rangle

\quark\_if\_recursion\_tail\_stop:on \* \quark\_if\_recursion\_tail\_stop:on

Tests if the \langle token list \rangle contains only \_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 insertion \rangle code is then added to the input stream after the recursion has ended.
Tests if (token list) contains only \q_recursion_tail, and if so terminates the recursion using \(\langle\text{type}\rangle\_\text{map\_break}\). The recursion end should be marked by \prg_break_point: Nn \(\langle\text{type}\rangle\_\text{map\_break}\).

18.4.1 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\ 3 built-in mapping functions, this mapping function cannot be nested, since the second map would overwrite the definition of \__my_map_dbl_fn:nn.
18.5 Scan marks

Scan marks are control sequences set equal to \texttt{\textbackslash 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 \texttt{l3regex}).

\begin{verbatim}
\scan_new:N \langle scan mark \rangle
\end{verbatim}

\begin{itemize}
  \item \texttt{\scan_new:N \langle scan mark \rangle}
  \end{itemize}

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

\begin{verbatim}
\s_stop
\end{verbatim}

\begin{itemize}
  \item \texttt{\s_stop}
  \end{itemize}

Used at the end of a set of instructions, as a marker that can be jumped to using \texttt{\use_-none\_delimit\_by\_s\_stop:w}.

\begin{verbatim}
\use_none\_delimit\_by\_s\_stop:w \star \use_none\_delimit\_by\_s\_stop:w \langle tokens \rangle \s_stop
\end{verbatim}

\begin{itemize}
  \item \texttt{\use_none\_delimit\_by\_s\_stop:w \star \use_none\_delimit\_by\_s\_stop:w \langle tokens \rangle \s_stop}
  \end{itemize}

Removes the \langle tokens \rangle and \texttt{\s_stop} from the input stream. This leads to a low-level \TeX{} error if \texttt{\s_stop} is absent.
Chapter 19

The \texttt{l3seq} package
Sequences and stacks

\texttt{l3seq} implements a “sequence” data type, which contain an ordered list of entries which may contain any 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 \texttt{l3seq}. This is achieved using a number of dedicated stack functions.

19.1 Creating and initialising sequences

\begin{Verbatim}
\texttt{\seq_new:N} \texttt{\seq_new:c} \texttt{\seq_set_eq:NN} \texttt{\seq_gset_eq:N} \texttt{\seq_gclear:N} \texttt{\seq_gclear:c}
\end{Verbatim}

\texttt{\seq_new:N} \texttt{\seq_new:c} \texttt{\seq_set_eq:NN} \texttt{\seq_gset_eq:N} \texttt{\seq_gclear:N} \texttt{\seq_gclear:c}

\begin{itemize}
\item \texttt{\seq_new:N} (sequence)
\end{itemize}

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

\begin{itemize}
\item \texttt{\seq_clear:N} (sequence)
\end{itemize}

Clears all items from the (sequence).

\begin{itemize}
\item \texttt{\seq_clear_new:N} (sequence)
\end{itemize}

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

\begin{itemize}
\item \texttt{\seq_set_eq:NN} (sequence$_1$) (sequence$_2$)
\end{itemize}

Sets the content of (sequence$_1$) equal to that of (sequence$_2$).
\seq_set_from_clist:NN \seq_set_from_clist:Nn \seq_set_from_clist:cn \seq_gset_from_clist:NN \seq_gset_from_clist:Nn

Converting the data in the \langle comma list \rangle into a \langle sequence \rangle: the original \langle comma list \rangle is unchanged.

\seq_const_from_clist:Nn \seq_const_from_clist:cn

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

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

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

\seq_set_split_keep_spaces:Nnn \seq_set_split_keep_spaces:NnV \seq_gset_split_keep_spaces:Nnn \seq_gset_split_keep_spaces:NnV

Splits the \langle token list \rangle into \langle items \rangle separated by \langle delimiter \rangle, and assigns the result to the \langle sequence \rangle. One set of outer braces is removed (if any) but any surrounding spaces are retained: any braces inside one or more spaces are therefore kept. Empty \langle items \rangle are preserved by \texttt{\seq_set_split_keep_spaces:Nnn}, and can be removed afterwards using \texttt{\seq_remove_all:Nn \langle sequence \rangle \{ \}}. The \langle delimiter \rangle may not contain \{, \} or \# (assuming \TeX's normal category code régime). If the \langle delimiter \rangle is empty, the \langle token list \rangle is split into \langle items \rangle as a \langle token list \rangle. See also \texttt{\seq_set_split:Nnn}, which removes spaces around the delimiters.

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

Concatenates the content of \langle sequence_2 \rangle and \langle sequence_3 \rangle together and saves the result in \langle sequence_1 \rangle. The items in \langle sequence_2 \rangle are placed at the left side of the new sequence.
Tests whether the \( \text{sequence} \) is currently defined. This does not check that the \( \text{sequence} \) really is a sequence variable.

19.2 Appending data to sequences

\begin{align*}
\text{\texttt{\seqputleft:Nn} \langle \text{sequence} \rangle \langle \text{item} \rangle} \\
\text{\texttt{\seqgputleft:Nn} \langle \text{sequence} \rangle \langle \text{item} \rangle}
\end{align*}

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

\begin{align*}
\text{\texttt{\seqputright:Nn} \langle \text{sequence} \rangle \langle \text{item} \rangle} \\
\text{\texttt{\seqgputright:Nn} \langle \text{sequence} \rangle \langle \text{item} \rangle}
\end{align*}

Appends the \( \langle \text{item} \rangle \) to the right of the \( \langle \text{sequence} \rangle \).

19.3 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, \textit{i.e.} setting the \( \langle \text{token list variable} \rangle \) used with \texttt{\tlset:Nn} and \textit{never} \texttt{\tlgset:Nn}.

\begin{align*}
\text{\texttt{\seqgetleft:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle} \\
\text{\texttt{\seqgetleft:cN}} \\
\text{\texttt{\seqgetleft:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle} \\
\text{\texttt{\seqgetleft:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle}
\end{align*}

Stores the left-most item from a \( \langle \text{sequence} \rangle \) in the \( \langle \text{token list variable} \rangle \) without removing it from the \( \langle \text{sequence} \rangle \). The \( \langle \text{token list variable} \rangle \) is assigned locally. If \( \langle \text{sequence} \rangle \) is empty the \( \langle \text{token list variable} \rangle \) is set to the special marker \texttt{\qno_value}.

\begin{align*}
\text{\texttt{\seqgetright:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle} \\
\text{\texttt{\seqgetright:cN}} \\
\text{\texttt{\seqgetright:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle} \\
\text{\texttt{\seqgetright:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle}
\end{align*}

Stores the right-most item from a \( \langle \text{sequence} \rangle \) in the \( \langle \text{token list variable} \rangle \) without removing it from the \( \langle \text{sequence} \rangle \). The \( \langle \text{token list variable} \rangle \) is assigned locally. If \( \langle \text{sequence} \rangle \) is empty the \( \langle \text{token list variable} \rangle \) is set to the special marker \texttt{\qno_value}.

\begin{align*}
\text{\texttt{\seqpopleft:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle} \\
\text{\texttt{\seqpopleft:cN}} \\
\text{\texttt{\seqpopleft:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle} \\
\text{\texttt{\seqpopleft:NN} \langle \text{sequence} \rangle \langle \text{token list variable} \rangle}
\end{align*}

Pops the left-most item from a \( \langle \text{sequence} \rangle \) into the \( \langle \text{token list variable} \rangle \), \textit{i.e.} removes the item from the sequence and stores it in the \( \langle \text{token list variable} \rangle \). Both of the variables are assigned locally. If \( \langle \text{sequence} \rangle \) is empty the \( \langle \text{token list variable} \rangle \) is set to the special marker \texttt{\qno_value}.
\seq_gpop_left:NN (sequence) (token list variable)

Pops the left-most item from a \textit{sequence} into the \textit{token list variable}, i.e. removes the item from the sequence and stores it in the \textit{token list variable}. The \textit{sequence} is modified globally, while the assignment of the \textit{token list variable} is local. If \textit{sequence} is empty the \textit{token list variable} is set to the special marker \texttt{\q_no_value}.

\seq_pop_right:NN (sequence) (token list variable)

Pops the right-most item from a \textit{sequence} into the \textit{token list variable}, i.e. removes the item from the sequence and stores it in the \textit{token list variable}. Both of the variables are assigned locally. If \textit{sequence} is empty the \textit{token list variable} is set to the special marker \texttt{\q_no_value}.

\seq_gpop_right:NN (sequence) (token list variable)

Pops the right-most item from a \textit{sequence} into the \textit{token list variable}, i.e. removes the item from the sequence and stores it in the \textit{token list variable}. The \textit{sequence} is modified globally, while the assignment of the \textit{token list variable} is local. If \textit{sequence} is empty the \textit{token list variable} is set to the special marker \texttt{\q_no_value}.

\seq_item:Nn \langle sequence \rangle \{ \langle integer expression \rangle \}

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

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

\seq_rand_item:N \langle seq var \rangle

Selects a pseudo-random item of the \textit{sequence}. If the \textit{sequence} is empty the result is empty. This is not available in older versions of \TeXX.

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

### 19.4 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.
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.

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.

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.

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.
19.5 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 (sequence)

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

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

\seq_remove_all:Nn (sequence) {⟨item⟩}

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

\seq_reverse:N (sequence)

Reverses the order of the items stored in the ⟨sequence⟩.

\seq_sort:Nn ⟨sequence⟩ {⟨comparison code⟩}

Sorts the items in the ⟨sequence⟩ according to the ⟨comparison code⟩, and assigns the result to ⟨sequence⟩. The details of sorting comparison are described in Section 6.1.

\seq_shuffle:N (seq var)

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

\TeXhackers 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 \sys_rand_seed: only has 28-bits. The use of \toks internally means that sequences with more than 32767 or 65535 items (depending on the engine) cannot be shuffled.

19.6 Sequence conditionals

\seq_if_empty_p:N ⟨sequence⟩

\seq_if_empty:NTF ⟨sequence⟩ {(true code)} {(false code)}

Tests if the ⟨sequence⟩ is empty (containing no items).
\seq_if_in:NnTF \seq_if_in:(NV|Nv|No|Nx|cV|cv|co|cx)TF

Tests if the \langle item\rangle is present in the \langle sequence\rangle.

19.7 Mapping over sequences

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.

\seq_map_function:NN \star
\seq_map_function:cN \star

\seq_map_variable:NNn \seq_map_variable:(Ncn|cNn|ccn)

Stores each \langle item\rangle of the \langle sequence\rangle in turn in the (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 item\rangle in the \langle sequence\rangle, or its original value if the \langle sequence\rangle is empty. The \langle items\rangle are returned from left to right.

\seq_map_indexed_function:NN \star

\seq_map_indexed_function:NN \langle sequence var\rangle \langle function\rangle

New: 2018-05-03

Applies \langle function\rangle to every entry in the \langle sequence variable\rangle. The \langle function\rangle should have signature :nn. It receives two arguments for each iteration: the \langle index\rangle (namely 1 for the first entry, then 2 and so on) and the \langle item\rangle.
\seq_map_indexed_inline:Nn (seq var) {(inline function)}

Applies \textit{(inline function)} to every entry in the \textit{(sequence variable)}. The \textit{(inline function)} should consist of code which receives the \textit{(index)} (namely 1 for the first entry, then 2 and so on) as \texttt{#1} and the \textit{(item)} as \texttt{#2}.

\seq_map_break: \star

Updated: 2012-06-29

\seq_map_break:\star

Used to terminate a $\backslash$seq_map\ldots function before all entries in the \textit{(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 $\backslash$seq_map\ldots scenario leads to low level TeX errors.

\textbf{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 \star

Updated: 2012-06-29

\seq_map_break:n {(code)}

Used to terminate a $\backslash$seq_map\ldots function before all entries in the \textit{(sequence)} have been processed, inserting the \textit{(code)} after the mapping has ended. This normally takes place within a conditional statement, for example

\begin{verbatim}
\seq_map_inline:Nn \l_my_seq
{\str_if_eq:nTF { #1 } { bingo }{ \seq_map_break:n { <code> } }
{ \% Do something useful}
}
\end{verbatim}

Use outside of a $\backslash$seq_map\ldots scenario leads to low level TeX errors.

\textbf{TeXhackers 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.
Applies \emph{(inline function)} to every \emph{(item)} stored within the \emph{(sequences)}. The \emph{(inline function)} should consist of code which will receive the \emph{(item)} as \#1. The sequence resulting applying \emph{(inline function)} to each \emph{(item)} is assigned to \emph{(sequences)}.

\textbf{\TeX{}hackers note:} Contrarily to other mapping functions, \texttt{\seq_set_map:NNn} cannot be used in this function, and would lead to low-level \TeX{} errors.

\seq_set_map_x:NNn \seq_gset_map_x:NNn

\textbf{\TeX{}hackers note:} Contrarily to other mapping functions, \texttt{\seq_set_map_break:} cannot be used in this function, and would lead to low-level \TeX{} errors.

\seq_count:N

Leaves the number of items in the \emph{(sequence)} in the input stream as an \emph{(integer denotation)}. The total number of items in a \emph{(sequence)} includes those which are empty and duplicates, \emph{i.e.} every item in a \emph{(sequence)} is unique.

\textbf{19.8 Using the content of sequences directly}

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

inserts “\texttt{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{\unexpanded} primitive \texttt{\exp_not:n}, which means that the \emph{(items)} do not expand further when appearing in an \texttt{x-type argument expansion}. 

148
\seq_use:Nn \seq_use:cn *

Places the contents of the \seq var in the input stream, with the \seq use \{ \textit{separator} \} between the items. If the sequence has a single item, it is placed in the input stream with no \seq use \{ \textit{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.

\TeXhacker\textbf{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.

### 19.9 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:cnF

If the \sequence is empty, leaves the \textit{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.
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 top item from the \langle sequence \rangle in the \langle token
list variable \rangle, i.e. removes the item from the \langle sequence \rangle. 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 top item from the \langle sequence \rangle in the \langle token
list variable \rangle, i.e. removes the item from the \langle sequence \rangle. The \langle sequence \rangle is modified globally,
while the \langle token list variable \rangle is assigned locally.

\seq_push:Nn \seq_push:(NV|Nv|No|Nx|cn|cV|cv|co|cx) \seq_push:Nn \seq_push:(NV|Nv|No|Nx|cn|cV|cv|co|cx)

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

19.10 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 \{ \langle item \rangle \}
tests if the \langle item \rangle is in the set.

Adding an \langle item \rangle 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 \{ \langle item \rangle \}
{ \seq_put_right:Nn \langle seq var \rangle \{ \langle item \rangle \} }

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

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

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 \seq_var_3
\seq_map_inline:Nn \seq_var_1
{
\seq_if_in:NnT \seq_var_2 {#1}
\seq_put_right:Nn \seq_var_3 {#1}
}

The code as written here only works if \seq_var_3 is different from the other two sequence variables. To cover all cases, items should first be collected in a sequence \l__\langle pkg\rangle\_internal_seq, then \seq_var_3 should be set equal to this internal sequence.
The same remark applies to other set functions.

The union of two sets \seq_var_1 and \seq_var_2 can be stored into \seq_var_3 through
\seq_concat:NNN \seq_var_3 \seq_var_1 \seq_var_2
\seq_remove_duplicates:N \seq_var_3
or by adding items to (a copy of) \seq_var_1 one by one
\seq_set_eq:NN \seq_var_3 \seq_var_1
\seq_map_inline:Nn \seq_var_2
{ \seq_if_in:NnF \seq_var_3 {#1}
\seq_put_right:Nn \seq_var_3 {#1} }

The second approach is faster than the first when the \seq_var_2 is short compared to \seq_var_1.

The difference of two sets \seq_var_1 and \seq_var_2 can be stored into \seq_var_3 by removing items of the \seq_var_2 from (a copy of) the \seq_var_1 one by one.
\seq_set_eq:NN \seq_var_3 \seq_var_1
\seq_map_inline:Nn \seq_var_2
{ \seq_remove_all:Nn \seq_var_3 {#1} }

The symmetric difference of two sets \seq_var_1 and \seq_var_2 can be stored into \seq_var_3 by computing the difference between \seq_var_1 and \seq_var_2 and storing the result as \l__\langle pkg\rangle\_internal_seq, then the difference between \seq_var_2 and \seq_var_1, and finally concatenating the two differences to get the symmetric differences.
\seq_set_eq:NN \seq_var_3 \l__\langle pkg\rangle\_internal_seq \seq_var_1
\seq_map_inline:Nn \seq_var_2
{ \seq_remove_all:Nn \seq_var_3 \l__\langle pkg\rangle\_internal_seq {#1} }
\seq_set_eq:NN \seq_var_3 \seq_var_2
\seq_map_inline:Nn \seq_var_1
{ \seq_remove_all:Nn \seq_var_3 {#1} }
\seq_concat:NNN \seq_var_3 \seq_var_3 \l__\langle pkg\rangle\_internal_seq

19.11 Constant and scratch sequences

\c_empty_seq
Constant that is always empty.

See: 2012-07-02
Scratch sequences 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.

\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 \LaTeX-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

19.12 Viewing sequences

\seq_show:N \seq_show:c
Updated: 2021-04-29

\seq_log:N \seq_log:c
New: 2014-08-12
Updated: 2021-04-29

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

\seq_log:N \{sequence\}
Writes the entries in the \{sequence\} in the log file.
Chapter 20

The 13int package
Integers

Calculation and comparison of integer values can be carried out using literal numbers, \texttt{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 ("\texttt{intexpr}").
20.1 Integer expressions

\int_eval:n \{⟨integer expression⟩\}

Evaluates the ⟨integer expression⟩ 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 ⟨integer expression⟩ should consist, after expansion, of \(+\), \(-\), \(*\), \(/\), \(\) 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 \(+\) or \(-\) at the start of an expression or after \(+\), \(-\), \(*\), \(/\) or \(\) 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 ⟨integer expression⟩ 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 ⟨internal integer⟩, 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.
\texttt{\int_eval:w} \texttt{\langle integer expression\rangle}

Evaluates the \texttt{\langle integer expression\rangle} as described for \texttt{\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 \texttt{\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, \texttt{\int_eval:w 1_{\texttt{+}}1_{\texttt{9}}} expands to 29 since the digit 9 is not part of the expression.

\texttt{\int_sign:n} \{\langle intexpr\rangle\}

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

\texttt{\int_abs:n} \{\langle intexpr\rangle\}

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

\texttt{\int_div_round:nn} \{\langle intexpr\rangle_{1}\} \{\langle intexpr\rangle_{2}\}

Evaluates the two \texttt{\langle 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 \texttt{/} directly in a \texttt{\langle integer expression\rangle}. The result is left in the input stream as an \texttt{\langle integer denotation\rangle} after two expansions.

\texttt{\int_div_truncate:nn} \{\langle intexpr\rangle_{1}\} \{\langle intexpr\rangle_{2}\}

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

\texttt{\int_max:nn} \{\langle intexpr\rangle_{1}\} \{\langle intexpr\rangle_{2}\}

\texttt{\int_min:nn} \{\langle intexpr\rangle_{1}\} \{\langle intexpr\rangle_{2}\}

Evaluates the \texttt{\langle integer expressions\rangle} as described for \texttt{\int_eval:n} and leaves either the larger or smaller value in the input stream as an \texttt{\langle integer denotation\rangle} after two expansions.

\texttt{\int_mod:nn} \{\langle intexpr\rangle_{1}\} \{\langle intexpr\rangle_{2}\}

Evaluates the two \texttt{\langle integer expressions\rangle} as described earlier, then calculates the integer remainder of dividing the first expression by the second. This is obtained by subtracting \texttt{\int_div_truncate:nn} \{\langle intexpr\rangle_{1}\} \{\langle intexpr\rangle_{2}\} times \texttt{\langle intexpr\rangle_{2}} from \texttt{\langle intexpr\rangle_{1}}. Thus, the result has the same sign as \texttt{\langle intexpr\rangle_{1}} and its absolute value is strictly less than that of \texttt{\langle intexpr\rangle_{2}}. The result is left in the input stream as an \texttt{\langle integer denotation\rangle} after two expansions.

20.2 Creating and initialising integers

\texttt{\int_new:N} \texttt{\langle integer\rangle}

Creates a new \texttt{\langle integer\rangle} or raises an error if the name is already taken. The declaration is global. The \texttt{\langle 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:TF *

New: 2012-03-03

20.3 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 add:Nn (integer) {\langle integer expression\rangle}

Adds the result of the \langle integer expression\rangle to the current content of the \langle integer\rangle.

\int const:Nn (integer) {\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 (integer)

Sets \langle integer\rangle to 0.

\int zero new:N (integer)

Ensures that the \langle integer\rangle exists globally by applying \int new:N if necessary, then applies \int (g)zero:N to leave the \langle integer\rangle set to zero.

\int set eq:NN (integer_1) (integer_2)

Sets the content of \langle integer_1\rangle equal to that of \langle integer_2\rangle.

\int if exist p:N (int)
\int if exist:NTF (int) {\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 decr:N (integer)

Decreases the value stored in \langle integer\rangle by 1.

\int incr:N (integer)

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

Subtracts the result of the \texttt{(integer expression)} from the current content of the \texttt{(integer)}.

\int_use:N \int_use:c

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}).

\TeX\hacksnote{\int_use:N is the \TeX\ primitive \texttt{\the}: this is one of several IS\TeX\ names for this primitive.}

\int_compare_p:nNn \int_compare:nNnTF

\int_compare_p:nNn {\langle integer expression1\rangle} {\langle relation\rangle} {\langle integer expression2\rangle}
\int_compare:nNnTF {\langle integer expression1\rangle} {\langle relation\rangle} {\langle integer expression2\rangle} {\langle true code\rangle} {\langle false code\rangle}

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)}:

\begin{itemize}
  \item Equal =
  \item Greater than >
  \item Less than <
\end{itemize}

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{⟨intexpr₁⟩} and \texttt{⟨intexpr₂⟩} using the \texttt{⟨relation₁⟩}, then \texttt{⟨intexpr₂⟩} and \texttt{⟨intexpr₃⟩} using the \texttt{⟨relation₂⟩}, until finally comparing \texttt{⟨intexpr₉⟩} and \texttt{⟨intexpr₉₊₁⟩} using the \texttt{⟨relation₉⟩}. 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 \hspace{2em} = \text{ or } ==
  \item Greater than or equal to \hspace{1em} \geq
  \item Greater than \hspace{2em} >
  \item Less than or equal to \hspace{1em} \leq
  \item Less than \hspace{2em} <
  \item Not equal \hspace{2em} 
\end{itemize}

This function is more flexible than \texttt{\textbackslash int\_compare:nNnTF} but around 5 times slower.
This function evaluates the \(test\ integer\ expression\) and compares this in turn to each of the \(integer\ expression\ cases\). If the two are equal 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 \int_case:nn, which does nothing if there is no match, is also available. For example

\[
\int_case:nn \left\{ 2 \times 5 \right\}
\{
\{ 5 \} \quad \{ Small \}
\{ 4 + 6 \} \quad \{ Medium \}
\{ -2 \times 10 \} \quad \{ Negative \}
\{ No idea! \}
\}
\]

leaves “Medium” in the input stream.

This function first evaluates the \(integer\ expression\) as described for \int_eval:n. It then evaluates if this is odd or even, as appropriate.

### 20.6 Integer expression loops

\[
\int_do_until:nNnn \left\{ \langle integer\ expression_1 \rangle \right\} \left\{ \langle relation \rangle \right\} \left\{ \langle integer\ expression_2 \rangle \right\} \left\{ \langle code \rangle \right\}
\]

Places the \(code\) in the input stream for \TeX\ to process, and then evaluates the relationship between the two \(integer\ expressions\) as described for \int_compare:nNnTF. If the test is \texttt{false} then the \(code\) is inserted into the input stream again and a loop occurs until the \(relation\) is \texttt{true}.

\[
\int_do_while:nNnn \left\{ \langle integer\ expression_1 \rangle \right\} \left\{ \langle relation \rangle \right\} \left\{ \langle integer\ expression_2 \rangle \right\} \left\{ \langle code \rangle \right\}
\]

Places the \(code\) in the input stream for \TeX\ to process, and then evaluates the relationship between the two \(integer\ expressions\) as described for \int_compare:nNnTF. If the test is \texttt{true} then the \(code\) is inserted into the input stream again and a loop occurs until the \(relation\) is \texttt{false}. 

159
\int_until_do:nNnn \{\text{relation}\} \{\text{code}\}

Evaluates the relationship between the two \textit{integer expressions} as described for \texttt{\int_compare:nNnTF}, and then places the \texttt{\texttt{code}} in the input stream if the \texttt{relation} is \texttt{false}. After the \texttt{\texttt{code}} has been processed by \LaTeX{} the test is repeated, and a loop occurs until the test is \texttt{true}.

\int_while_do:nNnn \{\text{relation}\} \{\text{code}\}

Evaluates the relationship between the two \textit{integer expressions} as described for \texttt{\int_compare:nNnTF}, and then places the \texttt{\texttt{code}} in the input stream if the \texttt{relation} is \texttt{true}. After the \texttt{\texttt{code}} has been processed by \LaTeX{} the test is repeated, and a loop occurs until the test is \texttt{false}.

\int_do_until:nn \{\text{integer relation}\} \{\text{code}\}

Places the \texttt{\texttt{code}} in the input stream for \LaTeX{} to process, and then evaluates the \texttt{integer relation} as described for \texttt{\int_compare:nNnTF}. If the test is \texttt{false} then the \texttt{\texttt{code}} is inserted into the input stream again and a loop occurs until the \texttt{relation} is \texttt{true}.

\int_do_while:nn \{\text{integer relation}\} \{\text{code}\}

Places the \texttt{\texttt{code}} in the input stream for \LaTeX{} to process, and then evaluates the \texttt{integer relation} as described for \texttt{\int_compare:nNnTF}. If the test is \texttt{true} then the \texttt{\texttt{code}} is inserted into the input stream again and a loop occurs until the \texttt{relation} is \texttt{false}.

\int_until_do:nn \{\text{integer relation}\} \{\text{code}\}

Evaluates the \texttt{integer relation} as described for \texttt{\int_compare:nNnTF}, and then places the \texttt{\texttt{code}} in the input stream if the \texttt{relation} is \texttt{false}. After the \texttt{\texttt{code}} has been processed by \LaTeX{} the test is repeated, and a loop occurs until the test is \texttt{true}.

\int_while_do:nn \{\text{integer relation}\} \{\text{code}\}

Evaluates the \texttt{\texttt{integer relation}} as described for \texttt{\int_compare:nNnTF}, and then places the \texttt{\texttt{code}} in the input stream if the \texttt{relation} is \texttt{true}. After the \texttt{\texttt{code}} has been processed by \LaTeX{} the test is repeated, and a loop occurs until the test is \texttt{false}. 

Updated: 2013-01-13
20.7 Integer step functions

\int_step_function:nN \{(final value)\} \{function\}
\int_step_function:nnN \{(initial value) \{final value\}\} \{function\}
\int_step_function:nnnN \{(initial value) \{step\} \{final value\}\} \{function\}

This function first evaluates the \{initial value\}, \{step\} and \{final value\}, all of which should be integer expressions. The \{function\} is then placed in front of each \{value\} from the \{initial value\} to the \{final value\} in turn (using \{step\} between each \{value\}). The \{step\} must be non-zero. If the \{step\} is positive, the loop stops when the \{value\} becomes larger than the \{final value\}. If the \{step\} is negative, the loop stops when the \{value\} becomes smaller than the \{final value\}. The \{function\} should absorb one numerical argument. For example

\cs_set:Npn \my_func:n #1 \{ [I saw #1] \quad \}
\int_step_function:nnnN \{ 1 \} \{ 1 \} \{ 5 \} \my_func:n

would print

[I saw 1] [I saw 2] [I saw 3] [I saw 4] [I saw 5]

The functions \int_step_function:nN and \int_step_function:nnN both use a fixed \{step\} of 1, and in the case of \int_step_function:nN the \{initial value\} is also fixed as 1. These functions are provided as simple short-cuts for code clarity.

\int_step_variable:nN \{(final value)\} \{tl var\} \{code\}
\int_step_variable:nnN \{(initial value) \{final value\}\} \{tl var\} \{code\}
\int_step_variable:nnnn \{(initial value) \{step\} \{final value\}\} \{tl var\} \{code\}

This function first evaluates the \{initial value\}, \{step\} and \{final value\}, all of which should be integer expressions. Then for each \{value\} from the \{initial value\} to the \{final value\} in turn (using \{step\} between each \{value\}), the \{code\} is inserted into the input stream with \#1 replaced by the current \{value\}. Thus the \{code\} should define a function of one argument (\#1).

The functions \int_step_variable:nN and \int_step_variable:nnN both use a fixed \{step\} of 1, and in the case of \int_step_variable:nN the \{initial value\} is also fixed as 1. These functions are provided as simple short-cuts for code clarity.
20.8 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 \{(integer expression)\} \{(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 \}
...}
26 \{ z \}
}
\int_to_bin:n *
\int_to_bin:n \{(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 *
\int_to_hex:n \{(integer expression)\}
\int_to_Hex:n *
\int_toHex:n \{(integer expression)\}
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 *
\int_to_oct:n \{(integer expression)\}
\int_to_Base:n *
\int_toBase:n \{(integer expression)\} \{(base)\}
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 *
\int_to_base:n \{(integer expression)\} \{(base)\}
\int_to_Base:nn *
\int_toBase:n \{(integer expression)\} \{(base)\}
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).

\textbf{\TeX}hackers note: This is a generic version of \texttt{\int_to_bin:n}, etc.

\int_to_roman:n ☆
\int_to_Roman:n ☆
\int_to_roman:n \{(integer expression)\}
\int_to_Roman:n \{(integer expression)\}
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{\~v} for 5000) is \textit{not} used. For instance \texttt{\int_to_roman:n \{ 8249 \}} expands to \texttt{mmmmmmcccxxl}x.

\textbf{20.9 Converting from other formats to integers}

\int_from_alph:n ☆
\int_from_alph:n \{(letters)\}
\int_fromAlph:n \{(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 \texttt{+} and \texttt{-}. This is the inverse function of \texttt{\int_to_alph:n} and \texttt{\int_to_Alph:n}.
\int_from_bin:n \*

**New:** 2014-02-11  
**Updated:** 2014-08-25

Converting the \langle binary number \rangle into the integer (base 10) representation and leaves this in the input stream. The \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 \int_to_bin:n.

\int_from_hex:n \*

**New:** 2014-02-11  
**Updated:** 2014-08-25

Converting the \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 \langle hexadecimal number \rangle by upper or lower case letters. The \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 \int_to_hex:n and \int_to_Hex:n.

\int_from_oct:n \*

**New:** 2014-02-11  
**Updated:** 2014-08-25

Converting the \langle octal number \rangle into the integer (base 10) representation and leaves this in the input stream. The \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 \int_to_oct:n.

\int_from_roman:n \*

**Updated:** 2014-08-25

Converting the \langle roman numeral \rangle into the integer (base 10) representation and leaves this in the input stream. The \langle roman numeral \rangle is first converted to a string, with no expansion. The \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 \int_to_roman:n and \int_to_Roman:n.

\int_from_base:nn \*

**Updated:** 2014-08-25

Converting the \langle number \rangle expressed in \langle base \rangle into the appropriate value in base 10. The \langle number \rangle is first converted to a string, with no expansion. The \langle number \rangle should consist of digits and letters (either lower or upper case), plus optionally a leading sign. The maximum \langle base \rangle value is 36. This is the inverse function of \int_to_base:nn and \int_to_Base:nn.

\int_rand:nn \*

**New:** 2016-12-06  
**Updated:** 2018-04-27

Evaluating the two \langle integer expressions \rangle and produces a pseudo-random number between the two (with bounds included). This is not available in older versions of Xe\LaTeX.

\int_rand:n \*

**New:** 2018-05-05

Evaluating the \langle integer expression \rangle then produces a pseudo-random number between 1 and the \langle integer \rangle (included). This is not available in older versions of Xe\LaTeX.

\section{Random integers}

164
20.11 Viewing integers

\texttt{\int\_show:N} \int\_show:N \langle \text{integer} \rangle

Displays the value of the \langle \text{integer} \rangle on the terminal.

\texttt{\int\_show:n} \int\_show:n \{\langle \text{integer expression} \rangle\}

Displays the result of evaluating the \langle \text{integer expression} \rangle on the terminal.

\texttt{\int\_log:N} \int\_log:N \langle \text{integer} \rangle

\texttt{\int\_log:n} \int\_log:n \{\langle \text{integer expression} \rangle\}

\texttt{\c\_zero\_int} \c\_zero\_int

\texttt{\c\_one\_int} \c\_one\_int

The maximum value that can be stored as an integer.

Maximum number of registers.

Maximum character code completely supported by the engine.

20.12 Constant integers

Integer values used with primitive tests and assignments: their self-terminating nature makes these more convenient and faster than literal numbers.

20.13 Scratch integers

Scratch integer for local assignment. These are never used by the kernel code, and so are safe for use with any \texttt{\LaTeX}\texttt{3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

Scratch integer for global assignment. These are never used by the kernel code, and so are safe for use with any \texttt{\LaTeX}\texttt{3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
20.14 Direct number expansion

\int_value:w \* \int_value:w \texttt{(integer)}
\int_value:w \texttt{(integer denotation)} \texttt{(optional space)}

Expands the following tokens until an \texttt{(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 \texttt{(integer)} can consist of any
number of signs (with intervening spaces) followed by

- an integer variable (in fact, any \TeX{} register except \texttt{\toks}) or
- explicit digits (or by '\texttt{(octal digits)} or "\texttt{(hexadecimal digits)} or '{\texttt{(character)}).

In this last case expansion stops once a non-digit is found; if that is a space it is removed
as in 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}.

20.15 Primitive conditionals

\if_int_compare:w \* \if_int_compare:w \texttt{(integer)} \texttt{(relation)} \texttt{(integer)}
\texttt{(true code)}
\else:
\texttt{(false code)}
\fi:

Compare two integers using \texttt{(relation)}, which must be one of \texttt{=} \texttt{,<} or \texttt{>} with category code
12. The \texttt{\else:} branch is optional.

\TeX{}hackers note: These are both names for the \TeX{} primitive \texttt{\ifnum}.

\if_case:w \* \if_case:w \texttt{(integer)} \texttt{(case)}
\texttt{(case)}
\or: \texttt{(case)}
\or: \ldots
\else: \texttt{(default)}
\fi:

Selects a case to execute based on the value of the \texttt{(integer)}. The first case \texttt{(case)} is executed if \texttt{(integer)} is 0, the second \texttt{(case)} if the \texttt{(integer)} is 1, etc. The \texttt{(integer)}
may be a literal, a constant or an integer expression (e.g. using \texttt{\int_eval:n}).

\TeX{}hackers note: These are the \TeX{} primitives \texttt{\ifcase} and \texttt{\or}.
\if_int_odd:w * \if_int_odd:w \langle tokens \rangle \langle optional space \rangle
\langle true code \rangle
\else:
\langle true code \rangle
\fi:

Expands \langle tokens \rangle until a non-numeric token or a space is found, and tests whether the resulting \langle integer \rangle is odd. If so, \langle true code \rangle is executed. The \else: branch is optional.

\TeXhackers note: This is the \TeX primitive \ifodd.
Chapter 21

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 \( \langle \text{height} \rangle \). In expansion-only contexts, a flag can only be “raised”: this increases the \( \langle \text{height} \rangle \) by 1. The \( \langle \text{height} \rangle \) 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 \( \langle \text{flag name} \rangle \) such as \texttt{str_missing}. The \( \langle \text{flag name} \rangle \) 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 \texttt{l3str-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.

21.1 Setting up flags

\begin{verbatim}
\flag_new:n \flag_new:n \{\langle \text{flag name} \rangle\}
\end{verbatim}

Creates a new flag with a name given by \( \langle \text{flag name} \rangle \), or raises an error if the name is already taken. The \( \langle \text{flag name} \rangle \) may not contain spaces. The declaration is global, but flags are always local variables. The \( \langle \text{flag} \rangle \) initially has zero height.

\begin{verbatim}
\flag_clear:n \flag_clear:n \{\langle \text{flag name} \rangle\}
\end{verbatim}

The \( \langle \text{flag} \rangle \)'s height is set to zero. The assignment is local.
\flag_clear_new:n \flag_clear_new:n {\langle flag name\rangle}
Ensures that the \langle flag \rangle exists globally by applying \flag_new:n if necessary, then applies \flag_clear:n, setting the height to zero locally.

\flag_show:n \flag_show:n {\langle flag name\rangle}
Displays the \langle flag \rangle’\s height in the terminal.

\flag_log:n \flag_log:n {\langle flag name\rangle}
Writes the \langle flag \rangle’\s height to the log file.

21.2 Expandable flag commands

\flag_if_exist:p:n \flag_if_exist:n {\langle flag name\rangle}
\flag_if_exist:n \TF \*  
This function returns \texttt{true} if the \langle flag name \rangle references a flag that has been defined previously, and \texttt{false} otherwise.

\flag_if_raised:p:n \flag_if_raised:n {\langle flag name\rangle}
\flag_if_raised:n \TF \*  
This function returns \texttt{true} if the \langle flag \rangle has non-zero height, and \texttt{false} if the \langle flag \rangle has zero height.

\flag_height:n \flag_height:n {\langle flag name\rangle}
Expands to the height of the \langle flag \rangle as an integer denotation.

\flag_raise:n \flag_raise:n {\langle flag name\rangle}
The \langle flag \rangle’\s height is increased by 1 locally.
Chapter 22

The \texttt{l3clist} package
Comma separated lists

Comma lists (in short, \texttt{clist}) 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 items separated by commas.

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\textbackslash d } \\
\clist_put_right:Nn \l_my_clist { -{e-}, , ,{f} }, \\
\end{verbatim}

results in \texttt{\l_my_clist} containing \texttt{a,b,c\textbackslash d,\textbackslash e-},\texttt{\{f\}} namely the five items \texttt{a, b, c\textbackslash d, e-} and \texttt{f}. Comma lists normally do not contain empty or blank items so the following gives an empty comma list:

\begin{verbatim}
\clist_clear_new:N \l_my_clist \\
\clist_set: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.

Any \texttt{n}-type token list is a valid comma list input for \texttt{l3clist} functions, which will split the token list at every comma and process the items as described above. On the other hand, \texttt{N}-type functions expect comma list variables, which are particular token list variables in which this processing of items (and removal of blank items) has already
occurred. Because comma list variables are token list variables, expanding them once
yields their items separated by commas, and \l_3\l_t functions such as \texttt{\tl_show:N} can be
applied to them. (These functions often have \l_3\texttt{clist} analogues, which should be preferred.)

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 \l_3\texttt{seq}) 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 pro-
duce errors when their argument contains \{, \} or \# (assuming the usual \TeX\ category
codes apply). The sequence data type should thus certainly be preferred to comma lists
to store such items.

22.1 Creating and initialising comma lists

\begin{verbatim}
\clist_new:N \clist_new:c
\clist_const:Nn \clist_const:(N|cn|cx)
New: 2014-07-05
\clist_clear:N \clist_clear:c \clist_gclear:N \clist_gclear:c
\clist_clear_new:N \clist_clear_new:c \clist_gclear_new:N \clist_gclear_new:c
\clist_set_eq:NN \clist_set_eq: (cN|Nc|cc)
\clist_gset_eq:NN \clist_gset_eq: (cN|Nc|cc)
\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
\end{verbatim}

Creates a new \langle comma list \rangle or raises an error if the name is already taken. The declaration
is global. The \langle comma list \rangle initially contains no items.

Creates a new constant \langle clist var \rangle or raises an error if the name is already taken. The
value of the \langle clist var \rangle is set globally to the \langle comma list \rangle.

Clears all items from the \langle comma list \rangle.

Ensures that the \langle comma list \rangle exists globally by applying \texttt{\clist_new:N} if necessary,
then applies \texttt{\clist_(g)clear:N} to leave the list empty.

Sets the content of \langle comma list_1 \rangle equal to that of \langle comma list_2 \rangle. To set a token list
variable equal to a comma list variable, use \texttt{\tl_set_eq:NN}. Conversely, setting a comma
list variable to a token list is unadvisable unless one checks space-trimming and related
issues.

Converts the data in the \langle sequence \rangle into a \langle comma list \rangle: the original \langle sequence \rangle 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

\clist_concat:NNN \langle \text{comma list}_1 \rangle \langle \text{comma list}_2 \rangle \langle \text{comma list}_3 \rangle

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

\clist_concat:NN \clist_concat:ccc 
\clist_gconcat:NN \clist_gconcat:ccc

\clist_if_exist_p:N \clist_if_exist_p:c \clist_if_exist:NTF \clist_if_exist:cF

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

Updated: 2011-09-05

22.2 Adding data to comma lists

\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 \langle \text{comma list} \rangle to contain the \langle \text{items} \rangle, 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 \langle \text{tokens} \rangle as a single \langle \text{item} \rangle even if the \langle \text{tokens} \rangle contain commas or spaces, add a set of braces: \clist_set:Nn \langle \text{comma list} \rangle \{ \langle \text{tokens} \rangle \}.

Updated: 2011-09-05

\clist_put_left:Nn \clist_put_left:(NV|No|Nx|cn|cV|co|cx) \clist_gput_left:Nn \clist_gput_left:(NV|No|Nx|cn|cV|co|cx)

Appends the \langle \text{items} \rangle to the left of the \langle \text{comma list} \rangle. 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 \langle \text{tokens} \rangle as a single \langle \text{item} \rangle even if the \langle \text{tokens} \rangle contain commas or spaces, add a set of braces: \clist_put_left:Nn \langle \text{comma list} \rangle \{ \langle \text{tokens} \rangle \}.

Updated: 2011-09-05

\clist_put_right:Nn \clist_put_right:(NV|No|Nx|cn|cV|co|cx) \clist_gput_right:Nn \clist_gput_right:(NV|No|Nx|cn|cV|co|cx)

Appends the \langle \text{items} \rangle to the right of the \langle \text{comma list} \rangle. 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 \langle \text{tokens} \rangle as a single \langle \text{item} \rangle even if the \langle \text{tokens} \rangle contain commas or spaces, add a set of braces: \clist_put_right:Nn \langle \text{comma list} \rangle \{ \langle \text{tokens} \rangle \}.
22.3 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 \langle comma list \rangle
\clist_remove_duplicates:c
\clist_gremove_duplicates:N
\clist_gremove_duplicates:c

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

\TeXhacks\TeXhackers note: This function iterates through every item in the \langle comma list \rangle and does a comparison with the \langle items \rangle already checked. It is therefore relatively slow with large comma lists. Furthermore, it may fail if any of the items in the \langle comma list \rangle contains \{, \}, or \# (assuming the usual \TeX category codes apply).

\clist_remove_all:Nn \clist_remove_all:cn \clist_gremove_all:Nn \clist_gremove_all:cn

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

\TeXhacks\TeXhackers note: The function may fail if the \langle item \rangle contains \{, \}, or \# (assuming the usual \TeX category codes apply).

\clist_reverse:N \clist_reverse:N \langle comma list \rangle
\clist_reverse:c
\clist_greverse:N
\clist_greverse:c

Reverses the order of items stored in the \langle comma list \rangle.

\TeXhacks\TeXhackers note: The result is returned within \unexpanded, which means that the comma list does not expand further when appearing in an \texttt{x}-type or \texttt{e}-type argument expansion.

\clist_sort:Nn \clist_sort:cn \clist_gsort:Nn \clist_gsort:cn

Sorts the items in the \langle clist var \rangle according to the \langle comparison code \rangle, and assigns the result to \langle clist var \rangle. The details of sorting comparison are described in Section 6.1.
### 22.4 Comma list conditionals

\begin{itemize}
\item \clist_if_empty_p:N
\item \clist_if_empty_p:c
\item \clist_if_empty:NTF
\end{itemize}

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

\begin{itemize}
\item \clist_if_empty_p:n
\item \clist_if_empty:nTF
\end{itemize}

Tests if the \textit{comma list} is empty (containing no items). The rules for space trimming are as for other \textit{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.

\begin{itemize}
\item \clist_if_in:NnTF
\item \clist_if_in:nnTF
\end{itemize}

Tests if the \textit{item} is present in the \textit{comma list}. In the case of an \textit{n}-type \textit{comma list}, the usual rules of space trimming and brace stripping apply. Hence,

\begin{verbatim}
\clist_if_in:nnTF { a , {b}~ , {b} , c } { b } {true} {false}
\end{verbatim}
yields \texttt{true}.

\texttt{T\TeX} hackers note: The function may fail if the \textit{item} contains \{, \}, or \# (assuming the usual \texttt{T\TeX} category codes apply).

### 22.5 Mapping over 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 \textit{function} or \textit{code} discussed below remain in effect after the loop.

When the comma list is given explicitly, as an \textit{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\textsubscript{u},u\}\{b\textsubscript{i},i\}\{c\}, then the arguments passed to the mapped function are ‘a’, ‘b\textsubscript{i}’, an empty argument, and ‘c’.

When the comma list is given as an \textit{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 \textit{n}-type comma lists.

\begin{itemize}
\item \clist_map_function:NN
\item \clist_map_function:cN
\item \clist_map_function:nN
\end{itemize}

Applies \textit{function} to every \textit{item} stored in the \textit{comma list}. The \textit{function} receives one argument for each iteration. The \textit{items} are returned from left to right. The function \clist_map_inline:Nn is in general more efficient than \clist_map_function:NN.
\clist_map_inline:Nn \clist_map_inline:cn \clist_map_inline:nn
\clist_map_variable:NNn \clist_map_variable:NNn \clist_map_variable:nNn
\clist_map_tokens:Nn \clist_map_tokens:nn \clist_map_tokens:cn
\clist_map_tokens:N \clist_map_tokens:cn \clist_map_tokens:nn
\clist_map_break: \clist_map_break:

\clist_map_inline:Nn \comma list \{ \inline function \}
Applies \inline function to every \item stored within the \comma list. The \inline function should consist of code which receives the \item as #1. The \items are returned from left to right.

\clist_map_variable:NNn \comma list \variable \{ \code \}
Stores each \item of the \comma list 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 \comma list, or its original value if there were no \item. The \items are returned from left to right.

\clist_map_tokens:Nn \clist_map_tokens:nn \clist_map_tokens:cn
\clist_map_tokens:N \clist_map_tokens:cn \clist_map_tokens:nn
\clist_map_break: \clist_map_break:

\clist_map_inline:Nn \l_my_clist
\str_if_eq:nnTF { #1 } { bingo }
\{ \clist_map_break: \%
\ Do something useful
\}

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 further items are taken from the input stream. This depends on the design of the mapping function.
\clist_map_break:n \{\langle\text{code}\rangle\}

Used to terminate a \clist_map\ldots function before all entries in the \langle\text{comma list}\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

\clist_map_inline:Nn \l_my_clist
{\str_if_eq:nnTF { #1 } { bingo }{ \clist_map_break:n { \langle\text{code}\rangle\} }% Do something useful \}

Use outside of a \clist_map\ldots 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.

\clist_count:N \langle\text{comma list}\rangle

Leaves the number of items in the \langle\text{comma list}\rangle in the input stream as an \langle\text{integer denotation}\rangle. The total number of items in a \langle\text{comma list}\rangle includes those which are duplicates, \textit{i.e.} every item in a \langle\text{comma list}\rangle is counted.

22.6 Using the content of comma lists directly

\clist_use:Nnnn \langle\text{clist var}\rangle {\langle\text{separator between two}\rangle} {\langle\text{separator between more than two}\rangle} {\langle\text{separator between final two}\rangle}

Places the contents of the \langle\text{clist var}\rangle in the input stream, with the appropriate \langle\text{separator}\rangle between the items. Namely, if the comma list 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 comma list has exactly two items, then they are placed in the input stream separated by the \langle\text{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_tmpa_clist { a , b , , c , {de} , f }
\clist_use:Nnnn \l_tmpa_clist { -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 comma list has more than 2 items.

\textbf{\TeX{}hackers note:} The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp_not:n}), which means that the \langle\text{items}\rangle do not expand further when appearing in an \texttt{x}-type argument expansion.
\textbf{\clist_use:Nn} \textbf{\clist_use:cn}

Places the contents of the \texttt{\clist var} in the input stream, with the \texttt{\separator} 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{\TeXhackers note:} The result is returned within the \texttt{\unexpanded} primitive (\texttt{\exp_not:n}), which means that the \texttt{\items} do not expand further when appearing in an \texttt{x}-type argument expansion.

\textbf{\clist_use:nnnn} \textbf{\clist_use:nn} \textbf{\clist_use:nnn} \textbf{\clist_use:nn}

Places the contents of the \texttt{\comma list} in the input stream, with the appropriate \texttt{\separator} between the items. As for \texttt{\clist_set:Nn}, 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. The \texttt{\separators} are then inserted in the same way as for \texttt{\clist_use:nnnn} and \texttt{\clist_use:nn}, respectively.

### 22.7 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.

\textbf{\clist_get:NN} \textbf{\clist_get:cN} \textbf{\clist_get:NN} \textbf{\clist_get:cN}

Stores the left-most item from a \texttt{\comma list} in the \texttt{\token list variable} without removing it from the \texttt{\comma list}. The \texttt{\token list variable} is assigned locally. In the non-branching version, if the \texttt{\comma list} is empty the \texttt{\token list variable} is set to the marker value \texttt{\q_no_value}.

\textbf{\clist_pop:NN} \textbf{\clist_pop:cN}

Pops the left-most item from a \texttt{\comma list} into the \texttt{\token list variable}, i.e. removes the item from the comma list and stores it in the \texttt{\token list variable}. Both of the variables are assigned locally.
\clist_gpop:NN \clist_gpop:CN

Pops the left-most item from a \langle comma list \rangle into the \langle token list variable \rangle, i.e. removes the item from the comma list and stores it in the \langle token list variable \rangle. The \langle comma list \rangle is modified globally, while the assignment of the \langle token list variable \rangle is local.

\clist_gpop:NN \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. Both the \langle comma list \rangle and the \langle token list variable \rangle are assigned locally.

\clist_gpush:NN \clist_gpush:CN

Adds the \langle \{items\} \rangle to the top of the \langle comma list \rangle. Spaces are removed from both sides of each item as for any n-type comma list.

22.8 Using a single item

\clist_item:Nn \clist_item:cn \clist_item:nn

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

\TeXhacksnote: 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.
\clist_rand_item:N \clist_rand_item:c \clist_rand_item:n

Selects a pseudo-random item of the (comma list). If the (comma list) has no item, the result is empty.

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

### 22.9 Viewing comma lists

\clist_show:N \clist_show:c

Displays the entries in the (comma list) in the terminal.

\clist_show:n \clist_show:n

Displays the entries in the comma list in the terminal.

\clist_log:N \clist_log:c

Waits the entries in the (comma list) in the log file. See also \texttt{\clist_show:N} which displays the result in the terminal.

\clist_log:n \clist_log:n

Waits the entries in the comma list in the log file. See also \texttt{\clist_show:n} which displays the result in the terminal.

### 22.10 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 \texttt{\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 \texttt{\LaTeX3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.
Chapter 23

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}, \texttt{\if\charcode:w}, and \texttt{\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 23.7.
23.1 Creating character tokens

\char_set_active_eq:NN \char_set_active_eq:Nc \char_gset_active_eq:NN \char_gset_active_eq:Nc

Sets the behaviour of the \textlangle char\textrangle in situations where it is active (category code 13) to be equivalent to that of the \textlangle function\textrangle. The category code of the \textlangle char\textrangle is unchanged by this process. The \textlangle function\textrangle may itself be an active character.

\char_set_active_eq:nN \char_set_active_eq:nc \char_gset_active_eq:nN \char_gset_active_eq:nc

Sets the behaviour of the \textlangle char\textrangle which has character code as given by the \textlangle integer expression\textrangle in situations where it is active (category code 13) to be equivalent to that of the \textlangle function\textrangle. The category code of the \textlangle char\textrangle is unchanged by this process. The \textlangle function\textrangle may itself be an active character.

\char_generate:nn \char_generate:nc

Generates a character token of the given \textlangle charcode\textrangle and \textlangle catcode\textrangle (both of which may be integer expressions). The \textlangle catcode\textrangle may be one of

- 1 (begin group)
- 2 (end group)
- 3 (math toggle)
- 4 (alignment)
- 6 (parameter)
- 7 (math superscript)
- 8 (math subscript)
- 11 (letter)
- 12 (other)
- 13 (active)

and other values raise an error. The \textlangle charcode\textrangle may be any one valid for the engine in use. Active characters cannot be generated in older versions of \TeX. Another way to build token lists with unusual category codes is \textbackslash regex\_replace:nnN \textlangle character\textrangle \{\textlangle replacement\textrangle\} \{tl var\}.

\textbf{\TeX hackers note:} Exactly two expansions are needed to produce the character.
Converts the ⟨char⟩ to the equivalent case-changed character as detailed by the function name (see \str_foldcase:n and \text_titlecase:n for details of these terms). The case mapping is carried out with no context-dependence (cf. \text_uppercase:n, etc.) The str versions always generate “other” (category code 12) characters, whilst the standard versions generate characters with the category code of the ⟨char⟩ (i.e. only the character code changes).

\char_lowercase:N \star \char_uppercase:N \star \char_titlecase:N \star \char_foldcase:N \star \char_str_lowercase:N \star \char_str_uppercase:N \star \char_str_titlecase:N \star \char_str_foldcase:N \star

\char_set_catcode_letter:N \langle \text{character} \rangle
\char_set_catcode_escape:N
\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

Sets the category code of the ⟨character⟩ to that indicated in the function name. Depending on the current category code of the ⟨token⟩ the escape token may also be needed:

\char_set_catcode_other:N \star

The assignment is local.
Sets the category code of the \textit{character} which has character code as given by the \textit{integer expression}. This version can be used to set up characters which cannot otherwise be given (cf. the \texttt{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 \texttt{\char_set_catcode_{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 { \texttt{\char\texttt{\char}} } { \texttt{\char\texttt{\char + 32}} }
\end{verbatim}

The setting applies within the current \TeX{} group.
\char_value_lccode:n \char_value_lccode:n \langle \text{integer expression} \rangle

Expands to the current lower case code of the \langle \text{character} \rangle with character code given by the \langle \text{integer expression} \rangle.

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

Displays the current lower case code of the \langle \text{character} \rangle with character code given by the \langle \text{integer expression} \rangle on the terminal.

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

\text{Updated: 2015-08-06}

Sets up the behaviour of the \langle \text{character} \rangle when found inside \text{\texttt{text\_uppercase:n}}\text{, such that} \langle \text{character}_1 \rangle \text{ will be converted into} \langle \text{character}_2 \rangle. \text{The two} \langle \text{characters} \rangle \text{ may be specified using an} \langle \text{integer expression} \rangle \text{ for the character code concerned. This may include the \TeX \ '\langle \text{character} \rangle' \text{ method for converting a single character into its character code:}

\begin{verbatim}
\char_set_uccode:nn \{ \char\{ a \} \} \{ \char\{ A \} \} \% Standard behaviour
\char_set_uccode:nn \{ \char\{ A \} \} \{ \char\{ A - 32 \} \}
\char_set_uccode:nn \{ 60 \} \{ 50 \}
\end{verbatim}

The setting applies within the current \TeX \ group.

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

\text{Updated: 2015-08-06}

Expands to the current upper case code of the \langle \text{character} \rangle with character code given by the \langle \text{integer expression} \rangle.

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

Displays the current upper case code of the \langle \text{character} \rangle with character code given by the \langle \text{integer expression} \rangle on the terminal.

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

\text{Updated: 2015-08-06}

This function sets up the math code of \langle \text{character} \rangle. \text{The} \langle \text{character} \rangle \text{ is specified as an} \langle \text{integer expression} \rangle \text{ which will be used as the character code of the relevant character. The setting applies within the current \TeX \ group.}

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

\text{Updated: 2015-08-06}

This function sets up the space factor for the \langle \text{character} \rangle. \text{The} \langle \text{character} \rangle \text{ is specified as an} \langle \text{integer expression} \rangle \text{ which will be used as the character code of the relevant character. The setting applies within the current \TeX \ group.}

184
\char_value.sfcode:n * \char_value.sfcode:n \{(integer expression)\}
Expands to the current space factor for the \textit{character} with character code given by the \textit{integer expression}.

\char_show.sfcode:n \char_show.sfcode:n \{(integer expression)\}
Displays the current space factor for the \textit{character} with character code given by the \textit{integer expression} on the terminal.

\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 \textit{active} (catcode 13). Each entry in the sequence consists of a single escaped token, for example \texttt{\textbackslash}. Active tokens should be added to the sequence when they are defined for general document use.

\l_char_special_seq
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 \textit{letter} (catcode 11) or \textit{other} (catcode 12). Each entry in the sequence consists of a single escaped token, for example \texttt{\textbackslash} for the backslash or \texttt{\{ for an opening brace. Escaped tokens should be added to the sequence when they are defined for general document use.

### 23.3 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.
23.4 Converting tokens

\texttt{\textbackslash token\_to\_meaning:N} \texttt{\textbackslash token\_to\_meaning:c}

Inserts the current meaning of the \texttt{token} into the input stream as a series of characters of category code 12 (other). This is the primitive \TeX description of the \texttt{token}, thus for example both functions defined by \texttt{\cs_set_nopar:Npn} and token list variables defined using \texttt{\tl_new:N} are described as macros.

\textbf{\TeX hackers note:} This is the \TeX primitive \texttt{\meaning}. The \texttt{token} 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 \texttt{N}-type arguments.

\texttt{\textbackslash token\_to\_str:N} \texttt{\textbackslash token\_to\_str:c}

Converts the given \texttt{token} into a series of characters with category code 12 (other). If the \texttt{token} is a control sequence, this will start with the current escape character with category code 12 (the escape character is part of the \texttt{token}). This function requires only a single expansion.

\textbf{\TeX hackers note:} \texttt{\token_to_str:N} is the \TeX primitive \texttt{\string} renamed. The \texttt{token} 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 \texttt{N}-type arguments.

23.5 Token conditionals

\texttt{\textbackslash token\_if\_group\_begin\_p:N} \texttt{\textbackslash token\_if\_group\_begin:p:N} \texttt{\textbackslash token\_if\_group\_begin:NTF} \texttt{\textbackslash token\_if\_group\_begin:NTF}

Tests if \texttt{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 \texttt{N}-type argument.

\texttt{\textbackslash token\_if\_group\_end\_p:N} \texttt{\textbackslash token\_if\_group\_end:p:N} \texttt{\textbackslash token\_if\_group\_end:NTF} \texttt{\textbackslash token\_if\_group\_end:NTF}

Tests if \texttt{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 \texttt{N}-type argument.

\texttt{\textbackslash token\_if\_math\_toggle\_p:N} \texttt{\textbackslash token\_if\_math\_toggle:p:N} \texttt{\textbackslash token\_if\_math\_toggle:NTF} \texttt{\textbackslash token\_if\_math\_toggle:NTF}

Tests if \texttt{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:NTF \token_if_alignment_p:N \token_if_alignment:NTF

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:NTF \token_if_parameter_p:N \token_if_parameter:NTF

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:NTF \token_if_math_superscript_p:N \token_if_math_superscript:NTF

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:NTF \token_if_math_subscript_p:N \token_if_math_subscript:NTF

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:NTF \token_if_space_p:N \token_if_space:NTF

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.

\token_if_letter_p:N \token_if_letter:NTF \token_if_letter_p:N \token_if_letter:NTF

Tests if \token has the category code of a letter token.

\token_if_other_p:N \token_if_other:NTF \token_if_other_p:N \token_if_other:NTF

Tests if \token has the category code of an “other” token.

\token_if_active_p:N \token_if_active:NTF \token_if_active_p:N \token_if_active:NTF

Tests if \token has the category code of an active character.

\token_if_eq_catcode_p:NN \token_if_eq_catcode:NNTF \token_if_eq_catcode_p:NN \token_if_eq_catcode:NNTF

Tests if the two \tokens have the same category code.

\token_if_eq_charcode_p:NN \token_if_eq_charcode:NNTF \token_if_eq_charcode_p:NN \token_if_eq_charcode:NNTF

Tests if the two \tokens have the same character code.
\token_if_eq_meaning_p:NN \token_if_eq_meaning_p:NN \token_if_eq_meaning_p:NN \token_if_eq_meaning_p:NN \token_if_eq_meaning:NNTF \token_if_eq_meaning:NNTF \token_if_eq_meaning:NNTF \token_if_eq_meaning:NNTF {\{true code\}} {\{false code\}}

Tests if the two \emph{tokens} have the same meaning when expanded.

\token_if_macro_p:N \token_if_macro_p:N \token_if_macro:NTF \token_if_macro:NTF \token_if_macro:NTF \token_if_macro:NTF \token_if_macro:NTF \token_if_macro:NTF \token_if_macro:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is a \TeX{} macro.

\token_if_cs_p:N \token_if_cs_p:N \token_if_cs:NTF \token_if_cs:NTF \token_if_cs:NTF \token_if_cs:NTF \token_if_cs:NTF \token_if_cs:NTF \token_if_cs:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is a control sequence.

\token_if_expandable_p:N \token_if_expandable_p:N \token_if_expandable:NTF \token_if_expandable:NTF \token_if_expandable:NTF \token_if_expandable:NTF \token_if_expandable:NTF \token_if_expandable:NTF \token_if_expandable:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is expandable. This test returns \emph{false} for an undefined token.

\token_if_long_macro_p:N \token_if_long_macro_p:N \token_if_long_macro:NTF \token_if_long_macro:NTF \token_if_long_macro:NTF \token_if_long_macro:NTF \token_if_long_macro:NTF \token_if_long_macro:NTF \token_if_long_macro:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is a long macro.

\token_if_protected_macro_p:N \token_if_protected_macro_p:N \token_if_protected_macro:NTF \token_if_protected_macro:NTF \token_if_protected_macro:NTF \token_if_protected_macro:NTF \token_if_protected_macro:NTF \token_if_protected_macro:NTF \token_if_protected_macro:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is a protected macro: for a macro which is both protected and long this returns \emph{false}.

\token_if_protected_long_macro_p:N \token_if_protected_long_macro_p:N \token_if_protected_long_macro:NTF \token_if_protected_long_macro:NTF \token_if_protected_long_macro:NTF \token_if_protected_long_macro:NTF \token_if_protected_long_macro:NTF \token_if_protected_long_macro:NTF \token_if_protected_long_macro:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is a protected long macro.

\token_if_chardef_p:N \token_if_chardef_p:N \token_if_chardef:NTF \token_if_chardef:NTF \token_if_chardef:NTF \token_if_chardef:NTF \token_if_chardef:NTF \token_if_chardef:NTF \token_if_chardef:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is defined to be a chardef.

\TeX{}hackers note: Booleans, boxes and small integer constants are implemented as \texttt{chardefs}.

\token_if_mathchardef_p:N \token_if_mathchardef_p:N \token_if_mathchardef_p:N \token_if_mathchardef_p:N \token_if_mathchardef:NTF \token_if_mathchardef:NTF \token_if_mathchardef:NTF \token_if_mathchardef:NTF {\{true code\}} {\{false code\}}

Tests if the \emph{token} is defined to be a mathchardef.
\token_if_font_selection_p:N \token_if_font_selection:NTF \token_if_font_selection_p:N \token_if_font_selection:NTF \token_if_font_selection_p:N \token_if_font_selection:NTF \token_if_font_selection_p:N \token_if_font_selection:NTF \token_if_font_selection_p:N \token_if_font_selection:NTF

Tests if the \langle token \rangle is defined to be a font selection command.

\token_if_dim_register_p:N \token_if_dim_register:NTF \token_if_dim_register_p:N \token_if_dim_register:NTF \token_if_dim_register_p:N \token_if_dim_register:NTF \token_if_dim_register_p:N \token_if_dim_register:NTF \token_if_dim_register_p:N \token_if_dim_register:NTF

Tests if the \langle token \rangle is defined to be a dimension register.

\token_if_int_register_p:N \token_if_int_register:NTF \token_if_int_register_p:N \token_if_int_register:NTF \token_if_int_register_p:N \token_if_int_register:NTF \token_if_int_register_p:N \token_if_int_register:NTF \token_if_int_register_p:N \token_if_int_register:NTF

Tests if the \langle token \rangle is defined to be an integer register.

\TeXhackers note: Constant integers may be implemented as integer registers, \texttt{\chardef}, or \texttt{\mathchardef} depending on their value.

\token_if_muskip_register_p:N \token_if_muskip_register:NTF \token_if_muskip_register_p:N \token_if_muskip_register:NTF \token_if_muskip_register_p:N \token_if_muskip_register:NTF \token_if_muskip_register_p:N \token_if_muskip_register:NTF \token_if_muskip_register_p:N \token_if_muskip_register:NTF

Tests if the \langle token \rangle is defined to be a muskip register.

\token_if_skip_register_p:N \token_if_skip_register:NTF \token_if_skip_register_p:N \token_if_skip_register:NTF \token_if_skip_register_p:N \token_if_skip_register:NTF \token_if_skip_register_p:N \token_if_skip_register:NTF \token_if_skip_register_p:N \token_if_skip_register:NTF

Tests if the \langle token \rangle is defined to be a skip register.

\token_if_toks_register_p:N \token_if_toks_register:NTF \token_if_toks_register_p:N \token_if_toks_register:NTF \token_if_toks_register_p:N \token_if_toks_register:NTF \token_if_toks_register_p:N \token_if_toks_register:NTF \token_if_toks_register_p:N \token_if_toks_register:NTF

Tests if the \langle token \rangle is defined to be a toks register (not used by \LaTeX{}3).

\token_if_primitive_p:N \token_if_primitive:NTF \token_if_primitive_p:N \token_if_primitive:NTF \token_if_primitive_p:N \token_if_primitive:NTF \token_if_primitive_p:N \token_if_primitive:NTF \token_if_primitive_p:N \token_if_primitive:NTF

Tests if the \langle token \rangle is an engine primitive. In Lua\TeX{} this includes primitive-like commands defined using \{token.set_lua\}.
This function compares the \textit{test token} in turn with each of the \textit{token cases}. If the two are equal (as described for \texttt{\token_if_eq_catcode:NNTF}, \texttt{\token_if_eq_charcode:NNTF}, and \texttt{\token_if_eq_meaning:NNTF}, respectively) 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 functions \texttt{\token_case_catcode:Nn}, \texttt{\token_case_charcode:Nn}, and \texttt{\token_case_meaning:Nn}, which do nothing if there is no match, are also available.

### 23.6 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 not skip spaces the predefined tests include both a space-respecting and space-skipping version. In addition, using \texttt{\peek_analysis_map_inline:n}, one can map through the following tokens in the input stream and repeatedly perform some tests.

\begin{verbatim}
\peek_after:Nw \peek_after:Nw \ token \ function \ (function) \ (token)
\peek_gafter:Nw \peek_gafter:Nw \ token \ function \ (function) \ (token)
\end{verbatim}

Locally sets the test variable \texttt{\l_peek_token} equal to \textit{token} (as an implicit token, not as a token list), and then expands the \textit{function}. The \textit{token} remains in the input stream as the next item after the \textit{function}. The \textit{token} here may be \texttt{\l}, \texttt{\{ or \}} (assuming normal \TeX{} category codes), i.e. it is not necessarily the next argument which would be grabbed by a normal function.

Globally sets the test variable \texttt{\g_peek_token} equal to \textit{token} (as an implicit token, not as a token list), and then expands the \textit{function}. The \textit{token} remains in the input stream as the next item after the \textit{function}. The \textit{token} here may be \texttt{\l}, \texttt{\{ or \}} (assuming normal \TeX{} category codes), 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.
\peek_catcode:NNTF \peek_catcode:NNTF {test token} {(true code)} {(false code)}

Tests if the next ⟨token⟩ in the input stream has the same category code as the ⟨test token⟩ (as defined by the test \token_if_eq_catcode: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).

\peek_catcode_ignore_spaces:NNTF \peek_catcode_ignore_spaces:NNTF {test token} {(true code)} {(false code)}

Tests if the next non-space ⟨token⟩ in the input stream has the same category code as the ⟨test token⟩ (as defined by the test \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 ⟨token⟩ is left in the input stream after the ⟨true code⟩ or ⟨false code⟩ (as appropriate to the result of the test).

\peek_catcode_remove:NNTF \peek_catcode_remove:NNTF {test token} {(true code)} {(false code)}

Tests if the next ⟨token⟩ in the input stream has the same category code as the ⟨test token⟩ (as defined by the test \token_if_eq_catcode: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_catcode_remove_ignore_spaces:NNTF \peek_catcode_remove_ignore_spaces:NNTF {test token} {(true code)} {(false code)}

Tests if the next non-space ⟨token⟩ in the input stream has the same category code as the ⟨test token⟩ (as defined by the test \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 ⟨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_charcode:NNTF \peek_charcode:NNTF {test token} {(true code)} {(false code)}

Tests if the next ⟨token⟩ in the input stream has the same character code as the ⟨test token⟩ (as defined by the test \token_if_eq_charcode: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).

\peek_charcode_ignore_spaces:NNTF \peek_charcode_ignore_spaces:NNTF {test token} {(true code)} {(false code)}

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 \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 left in the input stream after the ⟨true code⟩ or ⟨false code⟩ (as appropriate to the result of the test).
\peek_charcode_remove:NNTF \peek_charcode_remove:NNTF \peek_charcode_remove:NNTF \peek_charcode_remove:NNTF \peek_charcode_remove:NNTF (test token) \{\langle true code\rangle\} \{\langle false code\rangle\}

Tests if the next \langle token\rangle in the input stream has the same character code as the \langle test token\rangle (as defined by the test \token_if_eq_charcode:NNTF). Spaces are respected by the test and the \langle token\rangle is removed from the input stream if the test is true. The function then places either the \langle true code\rangle or \langle false code\rangle in the input stream (as appropriate to the result of the test).

\peek_charcode_remove_ignore_spaces:NNTF \peek_charcode_remove_ignore_spaces:NNTF \peek_charcode_remove_ignore_spaces:NNTF \peek_charcode_remove_ignore_spaces:NNTF \peek_charcode_remove_ignore_spaces:NNTF \peek_charcode_remove_ignore_spaces:NNTF (test token) \{\langle true code\rangle\} \{\langle false code\rangle\}

Tests if the next non-space \langle token\rangle in the input stream has the same character code as the \langle test token\rangle (as defined by the test \token_if_eq_charcode:NNTF). Spaces are respected by the test and the \langle token\rangle is removed from the input stream if the test is true. The function then places either the \langle true code\rangle or \langle false code\rangle in the input stream (as appropriate to the result of the test).

\peek_meaning:NNTF \peek_meaning:NNTF \peek_meaning:NNTF \peek_meaning:NNTF \peek_meaning:NNTF \peek_meaning:NNTF (test token) \{\langle true code\rangle\} \{\langle false code\rangle\}

Tests if the next \langle token\rangle in the input stream has the same meaning as the \langle test token\rangle (as defined by the test \token_if_eq_meaning:NNTF). Spaces are respected by the test and 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).

\peek_meaning_remove:NNTF \peek_meaning_remove:NNTF \peek_meaning_remove:NNTF \peek_meaning_remove:NNTF \peek_meaning_remove:NNTF \peek_meaning_remove:NNTF (test token) \{\langle true code\rangle\} \{\langle false code\rangle\}

Tests if the next \langle token\rangle in the input stream has the same meaning as the \langle test token\rangle (as defined by the test \token_if_eq_meaning:NNTF). Spaces are respected by the test and the \langle token\rangle is removed from the input stream if the test is true. The function then places either the \langle true code\rangle or \langle false code\rangle in the input stream (as appropriate to the result of the test).

\peek_meaning_remove_ignore_spaces:NNTF \peek_meaning_remove_ignore_spaces:NNTF \peek_meaning_remove_ignore_spaces:NNTF \peek_meaning_remove_ignore_spaces:NNTF \peek_meaning_remove_ignore_spaces:NNTF \peek_meaning_remove_ignore_spaces:NNTF (test token) \{\langle true code\rangle\} \{\langle false code\rangle\}

Tests if the next non-space \langle token\rangle in the input stream has the same meaning as the \langle test token\rangle (as defined by the test \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 \langle token\rangle is removed from the input stream if the test is true. The function then places either the \langle true code\rangle or \langle false code\rangle 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 \texttt{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 \LaTeX{}X3) 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 \texttt{N}-type argument. However, if the next \langle token \rangle is for instance \texttt{\c_space_token}, the test takes the \langle false \rangle branch, even though the next \langle token \rangle is in fact a valid \texttt{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).

\peek_analysis_map_inline:n \peek_analysis_map_inline:n \{(inline function)\}

Repeatedly removes one \langle token \rangle from the input stream and applies the \langle inline function \rangle to it, until \peek_analysis_map_break: is called. The \langle inline function \rangle receives three arguments for each \langle token \rangle in the input stream:

- \langle tokens \rangle, which both \texttt{o}-expand and \texttt{x}-expand to the \langle token \rangle. The detailed form of \langle tokens \rangle may change in later releases.
- \langle char code \rangle, a decimal representation of the character code of the \langle token \rangle, \(-1\) if it is a control sequence.
- \langle catcode \rangle, a capital hexadecimal digit which denotes the category code of the \langle token \rangle: 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). This can be converted to an integer by writing \texttt{\langle catcode \rangle}.

These arguments are the same as for \texttt{\tl_analysis_map_inline:nn} defined in \L3tl-analysis. The \langle char code \rangle and \langle catcode \rangle do not take the meaning of a control sequence or active character into account: for instance, upon encountering the token \texttt{\c_group_begin_token} in the input stream, \peek_analysis_map_inline:n calls the \langle inline function \rangle with \#1 being \texttt{\exp_not:n \{ \c_group_begin_token \}} (with the current implementation), \#2 being \(-1\), and \#3 being 0, as for any other control sequence. In contrast, upon encountering an explicit begin-group token \{, the \langle inline function \rangle is called with arguments \texttt{\exp_after:wN \{ \if_false: \fi: \}} \texttt{123} and 1.

The mapping is done at the current group level, \textit{i.e.} any local assignments made by the \langle inline function \rangle remain in effect after the loop. Within the code, \texttt{\l_peek_token} is set equal (as a token, not a token list) to the token under consideration.

\peek_analysis_map_break: \peek_analysis_map_break:n
\peek_analysis_map_break:n
\{(code)\}

Stops the \texttt{\peek_analysis_map_inline:n} loop from seeking more tokens, and inserts \langle code \rangle in the input stream (empty for \texttt{\peek_analysis_map_break}:).
\peek_regex:nTF \peek_regex:NTF

\peek_regex:nTF {\langle regex\rangle} {\langle true code\rangle} {\langle false code\rangle}

Tests if the \langle tokens\rangle that follow in the input stream match the \langle regular expression\rangle. Any \langle tokens\rangle that have been read are left in the input stream after the \langle true code\rangle or \langle false code\rangle (as appropriate to the result of the test). See \l3regex for documentation of the syntax of regular expressions. The \langle regular expression\rangle is implicitly anchored at the start, so for instance \peek_regex:nTF \{ a \} is essentially equivalent to \peek_charcode:nTF a.

\textbf{\TeXhacker{\text{Note:}}} Implicit character tokens are correctly considered by \peek_regex:nTF as control sequences, while functions that inspect individual tokens (for instance \peek_charcode:nTF) only take into account their meaning.

The \peek_regex:nTF function only inspects as few tokens as necessary to determine whether the regular expression matches. For instance \peek_regex:nTF \{ abc \| [a-z] \} \{ \} \{ \} abc will only inspect the first token a even though the first branch abc of the alternative is preferred in functions such as \peek_regex_remove_once:n. This may have an effect on tokenization if the input stream has not yet been tokenized and category codes are changed.

\peek_regex_remove_once:n \peek_regex_remove_once:N

\peek_regex_remove_once:nTF \peek_regex_remove_once:NTF

\peek_regex_remove_once:nTF {\langle regex\rangle} {\langle true code\rangle} {\langle false code\rangle}

Tests if the \langle tokens\rangle that follow in the input stream match the \langle regex\rangle. If the test is true, the \langle tokens\rangle are removed from the input stream and the \langle true code\rangle is inserted, while if the test is false, the \langle false code\rangle is inserted followed by the \langle tokens\rangle that were originally in the input stream. See \l3regex for documentation of the syntax of regular expressions. The \langle regular expression\rangle is implicitly anchored at the start, so for instance \peek_regex_remove_once:nTF \{ a \} is essentially equivalent to \peek_charcode_remove:NTF a.

\textbf{\TeXhacker{\text{Note:}}} Implicit character tokens are correctly considered by \peek_regex_remove_once:nTF as control sequences, while functions that inspect individual tokens (for instance \peek_charcode:NTF) only take into account their meaning.
If the ⟨tokens⟩ that follow in the input stream match the ⟨regex⟩, replaces them according to the ⟨replacement⟩ as for \regex_replace_once:nnN, and leaves the result in the input stream, after the ⟨true code⟩. Otherwise, leaves ⟨false code⟩ followed by the ⟨tokens⟩ that were originally in the input stream, with no modifications. See \bregex for documentation of the syntax of regular expressions and of the ⟨replacement⟩: for instance \0 in the ⟨replacement⟩ is replaced by the tokens that were matched in the input stream. The ⟨regular expression⟩ is implicitly anchored at the start. In contrast to \regex_replace_once:nnN, no error arises if the ⟨replacement⟩ leads to an unbalanced token list: the tokens are inserted into the input stream without issue.

\TeXhackers note: Implicit character tokens are correctly considered by \peek_regex_replace_once:nnTF as control sequences, while functions that inspect individual tokens (for instance \peek_charcode:NTF) only take into account their meaning.

### 23.7 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 Xe\TeX{} 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).

There are also a few internal tokens. The following list may be incomplete in some engines.

- Expanding \the/font results in a token that looks identical to the command that was used to select the current font (such as \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 \fi of a conditional is encountered before the conditional is evaluated.

- Expanding \noexpand ⟨token⟩ (when the ⟨token⟩ is expandable) results in an internal token, displayed (temporarily) as \notexpanded: ⟨token⟩, whose shape coincides with the ⟨token⟩ and whose meaning differs from \relax.
• An \texttt{outer} \texttt{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}.

• Some frozen tokens can only be accessed in interactive sessions: \texttt{\cr, \right, \endgroup, \fi, \inaccessible}.

• In \LaTeX{}, there is also the strange case of “bytes” \texttt{\texttt{\makeatletter\char1100\makeatother xy} where x, y are any two lowercase hexadecimal digits, so that the hexadecimal number ranges from \texttt{\texttt{\textbackslash text{110000}=11141128}} to \texttt{\texttt{\textbackslash text{1100ff}= 1114367}}. These are used to output individual bytes to files, rather than UTF-8. For the purposes of token comparisons they behave like non-expandable primitive control sequences (\textit{not characters}) whose \texttt{\meaning is \texttt{the\_character\_id} followed by the given byte. If this byte is in the range 80–ff this gives an “invalid utf-8 sequence” error: applying \texttt{\token\_to\_str:N} or \texttt{\token\_to\_meaning:N} to these tokens is unsafe. Unfortunately, they don’t seem to be detectable safely by any means except perhaps Lua code.

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 \texttt{explicit} character tokens. Category codes that a character token can have are listed below by giving a sample output of the \TeX{} primitive \texttt{\meaning}, together with their \LaTeX{}3 names and most common example:

1  begin-group character (\texttt{\texttt{\textbackslash group\_begin}, often \{}),
2  end-group character (\texttt{\texttt{\textbackslash group\_end}, often \}}),
3  math shift character (\texttt{\texttt{\textbackslash math\_toggle}, often \$}),
4  alignment tab character (\texttt{\texttt{\textbackslash alignment}, often \&}),
6  macro parameter character (\texttt{\texttt{\textbackslash parameter}, often \#}),
7  superscript character (\texttt{\texttt{\textbackslash math\_superscript}, often \textasciicircum}),
8  subscript character (\texttt{\texttt{\textbackslash math\_subscript}, often \_}),
10 blank space (\texttt{\texttt{\textbackslash space}, often character code 32}),
11 the letter (\texttt{\texttt{\textbackslash letter}, such as A}),
12 the character (\texttt{\texttt{\textbackslash other}, such as 0}).

Category code 13 (\texttt{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 \texttt{implicit} character tokens. The meaning is otherwise in the following list:

• a macro, used in \LaTeX{}3 for most functions and some variables (\texttt{tl, fp, seq, \ldots}),
• a primitive such as \texttt{\textbackslash def} or \texttt{\textbackslash topmark}, used in \LaTeX{}3 for some functions,
• a register such as \texttt{\textbackslash count123}, used in \LaTeX{}3 for the implementation of some variables (\texttt{int, dim, \ldots}).
• a constant integer such as \char"56 or \mathchar"121,
• a font selection command,
• undefined.

Macros can be \texttt{protected} or not, \texttt{long} or not (the opposite of what \LaTeX3 calls \texttt{nopar}), and \texttt{outer} or not (unused in \LaTeX3). Their \texttt{meaning} takes the form

\begin{verbatim}
(prefix) macro: (argument) \rightarrow (replacement)
\end{verbatim}

where \texttt{(prefix)} is among \texttt{protected\ long\ outer}, \texttt{(argument)} describes parameters that the macro expects, such as \texttt{#1#2#3}, and \texttt{(replacement)} describes how the parameters are manipulated, such as \texttt{\int_eval:n(#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 :N.

When a macro takes a delimited argument \TeX{} scans ahead until finding the delimiter (outside any pairs of begin-group/end-group explicit characters), and the resulting list of tokens (with outer braces removed) becomes the argument. Note that explicit space characters at the start of the argument are not ignored in this case (and they prevent brace-stripping).
Chapter 24

The \texttt{l3prop} package

Property lists

\LaTeX{}3 implements a "property list" data type, which contain an unordered list of entries each of which consists of a \texttt{⟨key⟩} and an associated \texttt{⟨value⟩}. The \texttt{⟨key⟩} and \texttt{⟨value⟩} may both be any \texttt{⟨balanced text⟩}, but the \texttt{⟨key⟩} is processed using \texttt{tl_to_str:n}, meaning that category codes are ignored. 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 \texttt{⟨key⟩}: if an entry is added to a property list which already contains the \texttt{⟨key⟩} then the new entry overwrites the existing one. The \texttt{⟨keys⟩} 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 \emph{input} parsed by the \texttt{l3keys} module.

### 24.1 Creating and initialising property lists

\texttt{\prop_new:N} \texttt{(⟨property list⟩)}

Creates a new \texttt{⟨property list⟩} or raises an error if the name is already taken. The declaration is global. The \texttt{⟨property list⟩} initially contains no entries.

\texttt{\prop_clear:N} \texttt{\prop_clear:c} \texttt{\prop_gclear:N} \texttt{\prop_gclear:c}

\texttt{\prop_clear:N} \texttt{⟨property list⟩}

Clears all entries from the \texttt{⟨property list⟩}.

\texttt{\prop_clear_new:N} \texttt{\prop_clear_new:c} \texttt{\prop_gclear_new:N} \texttt{\prop_gclear_new:c}

\texttt{\prop_clear_new:N} \texttt{⟨property list⟩}

Ensures that the \texttt{⟨property list⟩} exists globally by applying \texttt{\prop_new:N} if necessary, then applies \texttt{\prop_(g)clear:N} to leave the list empty.

198
\prop_set_eq:NN \prop_set_eq: \langle property list \rangle \langle property list \rangle

Sets the content of \langle property list \rangle equal to that of \langle property list \rangle.

\prop_gset_eq:NN \prop_gset_eq: \langle property list \rangle \langle property list \rangle

\prop_set_from_keyval:Nn \prop_set_from_keyval: \langle prop var \rangle \{
\langle key1 \rangle = \langle value1 \rangle ,
\langle key2 \rangle = \langle value2 \rangle , ...
\}

Sets \langle prop var \rangle to contain key–value pairs given in the second argument. If duplicate keys appear only one of the values is kept.

Spaces are trimmed around every \langle key \rangle and every \langle value \rangle, and if the result of trimming spaces consists of a single brace group then a set of outer braces is removed. This enables both the \langle key \rangle and the \langle value \rangle to contain spaces, commas or equal signs. The \langle key \rangle is then processed by \tl_to_str:n.

\prop_const_from_keyval:Nn \prop_const_from_keyval: \langle prop var \rangle \{
\langle key1 \rangle = \langle value1 \rangle ,
\langle key2 \rangle = \langle value2 \rangle , ...
\}

Creates a new constant \langle prop var \rangle or raises an error if the name is already taken. The \langle prop var \rangle is set globally to contain key–value pairs given in the second argument, processed in the way described for \prop_set_from_keyval:Nn. If duplicate keys appear only one of the values is kept.

24.2 Adding and updating property list entries

\prop_put:Nnn \prop_put:N \langle property list \rangle \langle key \rangle \langle value \rangle

\prop_gput:Nnn \prop_gput:N \langle property list \rangle \langle key \rangle \langle value \rangle

\prop_put_if_new:Nnn \prop_put_if_new: \langle property list \rangle \langle key \rangle \langle value \rangle

\prop_gput_if_new:Nnn \prop_gput_if_new: \langle property list \rangle \langle key \rangle \langle value \rangle

Adds an entry to the \langle property list \rangle which may be accessed using the \langle key \rangle and which has \langle value \rangle. 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. 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 present in the \langle property list \rangle then no action is taken. Otherwise, a new entry is added as described for \prop_put:Nnn.
Combines the key–value pairs of \( \langle \text{prop var}_2 \rangle \) and \( \langle \text{prop var}_3 \rangle \), and saves the result in \( \langle \text{prop var}_1 \rangle \). If a key appears in both \( \langle \text{prop var}_2 \rangle \) and \( \langle \text{prop var}_3 \rangle \) then the last value, namely the value in \( \langle \text{prop var}_3 \rangle \) is kept while the other is discarded.

Updates the \( \langle \text{prop var} \rangle \) by adding entries for each key–value pair given in the second argument. The addition is done through \prop_put:Nnn, hence if the \( \langle \text{prop var} \rangle \) already contains some of the keys, the corresponding values are discarded and replaced by those given in the key–value list. If duplicate keys appear in the key–value list then only one of the values is kept.

The function is equivalent to storing the key–value pairs in a temporary property variable using \prop_set_from_keyval:Nn, then combining \( \langle \text{prop var} \rangle \) with the temporary variable using \prop_concat:NNN. In particular, the \( \langle \text{keys} \rangle \) and \( \langle \text{values} \rangle \) are space-trimmed and unbraced as described in \prop_set_from_keyval:Nn.

### 24.3 Recovering values from property lists

Recover the \( \langle \text{value} \rangle \) stored with \( \langle \text{key} \rangle \) from the \( \langle \text{property list} \rangle \), and places this in the \( \langle \text{token list variable} \rangle \). If the \( \langle \text{key} \rangle \) is not found in the \( \langle \text{property list} \rangle \) then the \( \langle \text{token list variable} \rangle \) is set to the special marker \q_no_value. The \( \langle \text{token list variable} \rangle \) is set within the current \TeX{} group. See also \prop_get:NnNTF.

Recover the \( \langle \text{value} \rangle \) stored with \( \langle \text{key} \rangle \) from the \( \langle \text{property list} \rangle \), and places this in the \( \langle \text{token list variable} \rangle \). If the \( \langle \text{key} \rangle \) is not found in the \( \langle \text{property list} \rangle \) then the \( \langle \text{token list variable} \rangle \) is set to the special marker \q_no_value. The \( \langle \text{key} \rangle \) and \( \langle \text{value} \rangle \) are then deleted from the property list. Both assignments are local. See also \prop_pop:NnNTF.

Recover the \( \langle \text{value} \rangle \) stored with \( \langle \text{key} \rangle \) from the \( \langle \text{property list} \rangle \), and places this in the \( \langle \text{token list variable} \rangle \). If the \( \langle \text{key} \rangle \) is not found in the \( \langle \text{property list} \rangle \) then the \( \langle \text{token list variable} \rangle \) is set to the special marker \q_no_value. The \( \langle \text{key} \rangle \) and \( \langle \text{value} \rangle \) are then deleted from the property list. The \( \langle \text{property list} \rangle \) is modified globally, while the assignment of the \( \langle \text{token list variable} \rangle \) is local. See also \prop_gpop:NnNTF.
\prop_item:Nn \prop_item:cn

Expands to the \textit{value} corresponding to the \textit{key} in the \textit{property list}. If the \textit{key} is missing, this has an empty expansion.

\TeXhacksnote{This function is slower than the non-expandable analogue \prop_get:Nn. The result is returned within the \texttt{unexpanded} primitive (\texttt{exp_not:n}), which means that the \textit{value} does not expand further when appearing in an \texttt{x}-type argument expansion.}

\prop_count:N
\prop_count:c

Leaves the number of key–value pairs in the \textit{property list} in the input stream as an \textit{integer denotation}.

24.4 Modifying property lists

\prop_remove:Nn \prop_remove: \prop_gremove:Nn \prop_gremove:

Removes the entry listed under \textit{key} from the \textit{property list}. If the \textit{key} is not found in the \textit{property list} no change occurs, \textit{i.e} there is no need to test for the existence of a key before deleting it.

24.5 Property list conditionals

\prop_if_exist_p:N \prop_if_exist:NTF \prop_if_exist:N \prop_if_exist:p:N
\prop_if_empty_p:N \prop_if_empty:NTF \prop_if_empty:N \prop_if_empty:p:N
\prop_if_in_p:N \prop_if_in:NTF \prop_if_in:N \prop_if_in:p:N

Tests whether the \textit{property list} is currently defined. This does not check that the \textit{property list} really is a property list variable.

Tests if the \textit{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{str_if_eq:nnTF}.

\TeXhacksnote{This function iterates through every key–value pair in the \textit{property list} and is therefore slower than using the non-expandable \prop_get:NnTF.}
24.6 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 value. This makes them useful for cases where different cases depend on the presence or absence of a key in a property list. They offer increased readability and performance over separate testing and recovery phases.

\prop_get:NnNTF \langle property list \rangle \{\langle key \rangle \} \langle \langle token list variable \rangle \rangle \{\langle true code \rangle \} \{\langle false code \rangle \}

If the \langle key \rangle is not present in the \langle property list \rangle, 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 key \rangle is present in the \langle property list \rangle, stores the corresponding \langle value \rangle in the \langle token list variable \rangle without removing it from the \langle property list \rangle, then leaves the \langle true code \rangle in the input stream. The \langle token list variable \rangle is assigned locally.

\prop_pop:NnNTF \langle property list \rangle \{\langle key \rangle \} \langle \langle token list variable \rangle \rangle \{\langle true code \rangle \} \{\langle false code \rangle \}

If the \langle key \rangle is not present in the \langle property list \rangle, 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 key \rangle is present in the \langle property list \rangle, pops the corresponding \langle value \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle property list \rangle. Both the \langle property list \rangle and the \langle token list variable \rangle are assigned locally.

\prop_gpop:NnNTF \langle property list \rangle \{\langle key \rangle \} \langle \langle token list variable \rangle \rangle \{\langle true code \rangle \} \{\langle false code \rangle \}

If the \langle key \rangle is not present in the \langle property list \rangle, 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 key \rangle is present in the \langle property list \rangle, pops the corresponding \langle value \rangle in the \langle token list variable \rangle, i.e. removes the item from the \langle property list \rangle. The \langle property list \rangle is modified globally, while the \langle token list variable \rangle is assigned locally.

24.7 Mapping over property lists

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

\prop_map_function:NN \langle property list \rangle \langle function \rangle

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.
\prop_map_inline:Nn\prop_map_inline:cn

Updated: 2013-01-08

\prop_map_tokens:Nn\prop_map_tokens:cn

Updated: 2013-01-08

\prop_map_break:

Updated: 2012-06-29

\prop_map_inline:Nn \l_my_prop
{ \str_if_eq:nnTF { #1 } { bingo } { \prop_map_break: } { % Do something useful } }

Use outside of a \prop_map\ldots scenario leads to low level TpX errors.

\textbf{TpXhackers 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 \{\texttt{code}\}

Used to terminate a \prop_map\ldots function before all entries in the \textit{property list} have been processed, inserting the \texttt{\textit{code}} after the mapping has ended. 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:n { <code> }
}{%
% 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 the \texttt{\textit{code}} is inserted into the input stream. This depends on the design of the mapping function.

### 24.8 Viewing property lists

\begin{verbatim}
\prop_show:N \prop_show:c
\end{verbatim}

Displays the entries in the \textit{property list} in the terminal.

\begin{verbatim}
\prop_log:N \prop_log:c
\end{verbatim}

Writes the entries in the \textit{property list} in the log file.

### 24.9 Scratch property lists

\begin{verbatim}
\l_tmpa_prop \l_tmpb_prop
\end{verbatim}

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.

\begin{verbatim}
\g_tmpa_prop \g_tmpb_prop
\end{verbatim}

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.
24.10 Constants

\texttt{\textbackslash c\_empty\_prop} A permanently-empty property list used for internal comparisons.
Chapter 25

The \texttt{l3skip} package
Dimensions and skips

\LaTeX{} 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.

25.1 Creating and initialising \texttt{dim} variables

\begin{verbatim}
\dim_new:N \dim_new:c
\dim_new:N \dim_new:c \dim_new:cn
\dim_zero:N \dim_zero:c \dim_gzero:N \dim_gzero:c
\dim_zero_new:N \dim_zero_new:c \dim_gzero_new:N \dim_gzero_new:c
\end{verbatim}

\begin{verbatim}
\dim_new:N \dim_new:c \dim_new:cn
\dim_new:N \dim_new:cn \dim_new:cn
\dim_zero:N \dim_zero:c \dim_gzero:N \dim_gzero:c
\dim_zero_new:N \dim_zero_new:c \dim_gzero_new:N \dim_gzero_new:c
\end{verbatim}

\texttt{\dim_new:N \dim_new:c \dim_new:cn}

Creates a new \texttt{(dimension)} or raises an error if the name is already taken. The declaration is global. The \texttt{(dimension)} is initially equal to \texttt{0 pt}.

\texttt{\dim_zero:N \dim_zero:c \dim_gzero:N \dim_gzero:c}

Sets \texttt{(dimension)} to \texttt{0 pt}.

\texttt{\dim_zero_new:N \dim_zero_new:c \dim_gzero_new:N \dim_gzero_new:c}

Ensures that the \texttt{(dimension)} exists globally by applying \texttt{\dim_new:N} if necessary, then applies \texttt{\dim_(g)zero:N} to leave the \texttt{(dimension)} set to zero.
Tests whether the \texttt{dimension} is currently defined. This does not check that the \texttt{dimension} really is a dimension variable.

### 25.2 Setting \texttt{dim} variables

\begin{verbatim}
\texttt{dim_add:Nn} \ (\texttt{dimension}) \ \{(\texttt{dimension expression})\}
\end{verbatim}

Adds the result of the \texttt{dimension expression} to the current content of the \texttt{dimension}.

\begin{verbatim}
\texttt{dim_set:Nn} \ (\texttt{dimension}) \ \{(\texttt{dimension expression})\}
\end{verbatim}

Sets \texttt{dimension} to the value of \texttt{dimension expression}, which must evaluate to a length with units.

\begin{verbatim}
\texttt{dim_set_eq:NN} \ (\texttt{dimension}_1) \ (\texttt{dimension}_2)
\end{verbatim}

Sets the content of \texttt{dimension}_1 equal to that of \texttt{dimension}_2.

\begin{verbatim}
\texttt{dim_sub:Nn} \ (\texttt{dimension}) \ \{(\texttt{dimension expression})\}
\end{verbatim}

Subtracts the result of the \texttt{dimension expression} from the current content of the \texttt{dimension}.

### 25.3 Utilities for \texttt{dimension} calculations

\begin{verbatim}
\texttt{dim_abs:n} \ (\texttt{dimexpr})
\end{verbatim}

Converts the \texttt{dimexpr} to its absolute value, leaving the result in the input stream as a \texttt{dimension denotation}.

\begin{verbatim}
\texttt{dim_max:nn} \ (\texttt{dimexpr}_1) \ (\texttt{dimexpr}_2)
\end{verbatim}

\begin{verbatim}
\texttt{dim_min:nn} \ (\texttt{dimexpr}_1) \ (\texttt{dimexpr}_2)
\end{verbatim}

Evaluates the two \texttt{dimension expressions} and leaves either the maximum or minimum value in the input stream as appropriate, as a \texttt{dimension denotation}.
\dim_ratio:nn \{dimexpr_1\} \{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 \text{ pt} * \dim_ratio:nn \{ 5 \text{ pt} \} \{ 10 \text{ 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 \text{ pt} \} \{ 10 \text{ pt} \} }
\tl_show:N \l_my_tl
\end{verbatim}

displays \texttt{327680/655360} on the terminal.

25.4 Dimension expression conditionals

\dim_compare_p:nNn \{dimexpr_1\} \{relation\} \{dimexpr_2\}
\dim_compare:nNnTF
\{dimexpr_1\} \{relation\} \{dimexpr_2\}
\{(true code)\} \{(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}:

Equal =
Greater than >
Less than <

This function is less flexible than \texttt{\dim_compare:nTF} but around 5 times faster.
\dim_compare_p:n \dim_compare:nTF

\dim_compare_p:n
\dim_compare:nTF

{ \langle \mbox{dimexpr} 1 \rangle \langle \mbox{relation} 1 \rangle 
\ldots
\langle \mbox{dimexpr} N \rangle \langle \mbox{relation} N \rangle
\langle \mbox{dimexpr} N+1 \rangle
}

\dim_compare:nTF
{ \langle \mbox{dimexpr} 1 \rangle \langle \mbox{relation} 1 \rangle 
\ldots
\langle \mbox{dimexpr} N \rangle \langle \mbox{relation} N \rangle
\langle \mbox{dimexpr} N+1 \rangle
}
{\langle \mbox{true code} \rangle} {\langle \mbox{false code} \rangle}

This function evaluates the \emph{(dimension expressions)} as described for \dim_eval:n and compares consecutive result using the corresponding \emph{(relation)}, namely it compares \langle \mbox{dimexpr} 1 \rangle and \langle \mbox{dimexpr} 2 \rangle using the \langle \mbox{relation} 1 \rangle, then \langle \mbox{dimexpr} 2 \rangle and \langle \mbox{dimexpr} 3 \rangle using the \langle \mbox{relation} 2 \rangle, until finally comparing \langle \mbox{dimexpr} N \rangle and \langle \mbox{dimexpr} N+1 \rangle using the \langle \mbox{relation} N \rangle. The test yields \texttt{true} if all comparisons are \texttt{true}. Each \emph{(dimension expression)} is evaluated only once, and the evaluation is lazy, in the sense that if one comparison is \texttt{false}, then no other \emph{(dimension expression)} is evaluated and no other comparison is performed. The \emph{(relations)} can be any of the following:

- Equal $=$ or $==$
- Greater than or equal to $\geq$
- Greater than $>$
- Less than or equal to $\leq$
- Less than $<$
- Not equal $!=$

This function is more flexible than \dim_compare:nTF but around 5 times slower.
\dim_case:nn \dim_case:nnTF \{[test\ dimension\ expression]\}\{
\{\dimexpr\ \case_1\}\{\code_1\}\}
\{\dimexpr\ \case_2\}\{\code_2\}\}
\ldots
\{\dimexpr\ \case_n\}\{\code_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.

25.5 Dimension expression loops

\dim_do_until:nNnn \dim_do_until:nNnn \{[expr_1]\} \{[relation]\} \{[expr_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 \{[expr_1]\} \{[relation]\} \{[expr_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 \{[expr_1]\} \{[relation]\} \{[expr_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}.
\texttt{\textbackslash dim\_while\_do:nNnn} \expiry\ \texttt{\textbackslash dim\_while\_do:nNnn \{\textbackslash \textit{dimexpr}_1\} \{\textbackslash \textit{relation}\} \{\textbackslash \textit{dimexpr}_2\} \{\textbackslash \textit{code}\}}

Evaluates the relationship between the two \textit{dimension expressions} as described for \texttt{\textbackslash dim\_compare:nNTF}, and then places the \textit{code} in the input stream if the \textit{relation} is true. After the \textit{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is false.

\texttt{\textbackslash dim\_do\_until:nn} \expiry\ \texttt{\textbackslash dim\_do\_until:nn \{\textbackslash \textit{dimension relation}\} \{\textbackslash \textit{code}\}}

Places the \textit{code} in the input stream for \TeX{} to process, and then evaluates the \textit{dimension relation} as described for \texttt{\textbackslash dim\_compare:nTF}. If the test is false then the \textit{code} is inserted into the input stream again and a loop occurs until the \textit{relation} is true.

\texttt{\textbackslash dim\_while\_do:nn} \expiry\ \texttt{\textbackslash dim\_while\_do:nn \{\textbackslash \textit{dimension relation}\} \{\textbackslash \textit{code}\}}

Places the \textit{code} in the input stream for \TeX{} to process, and then evaluates the \textit{dimension relation} as described for \texttt{\textbackslash dim\_compare:nTF}. If the test is true then the \textit{code} is inserted into the input stream again and a loop occurs until the \textit{relation} is false.

\texttt{\textbackslash until\_do:nn} \expiry\ \texttt{\textbackslash until\_do:nn \{\textbackslash \textit{dimension relation}\} \{\textbackslash \textit{code}\}}

Evaluates the \textit{dimension relation} as described for \texttt{\textbackslash dim\_compare:nTF}, and then places the \textit{code} in the input stream if the \textit{relation} is false. After the \textit{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is true.

\texttt{\textbackslash dim\_while\_do:nn} \expiry\ \texttt{\textbackslash dim\_while\_do:nn \{\textbackslash \textit{dimension relation}\} \{\textbackslash \textit{code}\}}

Evaluates the \textit{dimension relation} as described for \texttt{\textbackslash dim\_compare:nTF}, and then places the \textit{code} in the input stream if the \textit{relation} is true. After the \textit{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is false.

### 25.6 Dimension step functions

\texttt{\textbackslash \textit{dimstep}\_function:nNN} \expiry\ \texttt{\textbackslash \textit{dimstep}\_function:nNN \{\textbackslash \textit{initial value}\} \{\textbackslash \textit{step}\} \{\textbackslash \textit{final value}\} \{\textbackslash \textit{function}\}}

This function first evaluates the \textit{initial value}, \textit{step} and \textit{final value}, all of which should be dimension expressions. The \textit{function} is then placed in front of each \textit{value} from the \textit{initial value} to the \textit{final value} in turn (using \textit{step} between each \textit{value}). The \textit{step} must be non-zero. If the \textit{step} is positive, the loop stops when the \textit{value} becomes larger than the \textit{final value}. If the \textit{step} is negative, the loop stops when the \textit{value} becomes smaller than the \textit{final value}. The \textit{function} should absorb one argument.

\texttt{\textbackslash \textit{dimstep}\_inline:nnn} \expiry\ \texttt{\textbackslash \textit{dimstep}\_inline:nnn \{\textbackslash \textit{initial value}\} \{\textbackslash \textit{step}\} \{\textbackslash \textit{final value}\} \{\textbackslash \textit{code}\}}

This function first evaluates the \textit{initial value}, \textit{step} and \textit{final value}, all of which should be dimension expressions. Then for each \textit{value} from the \textit{initial value} to the \textit{final value} in turn (using \textit{step} between each \textit{value}), the \textit{code} is inserted into the input stream with \texttt{#1} replaced by the current \textit{value}. Thus the \textit{code} should define a function of one argument ($#1$).
This function first evaluates the ⟨initial value⟩, ⟨step⟩ and ⟨final value⟩, all of which should be dimension expressions. Then for each ⟨value⟩ from the ⟨initial value⟩ to the ⟨final value⟩ in turn (using ⟨step⟩ between each ⟨value⟩), the ⟨code⟩ is inserted into the input stream, with the ⟨tl var⟩ defined as the current ⟨value⟩. Thus the ⟨code⟩ should make use of the ⟨tl var⟩.

25.7 Using dim expressions and variables

\dim_sign:n \{ \dimexpr \}

Evaluates the ⟨dimexpr⟩ then leaves 1 or 0 or −1 in the input stream according to the sign of the result.

\dim_use:N \dim_use:c

Recovers the content of a ⟨dimension⟩ 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 ⟨dimension⟩ is required (such as in the argument of \dim_eval:n).

\TeXhacks{\dim_use:N is the \TeX primitive \the: this is one of several \LaTeX names for this primitive.}

\dim_to_decimal:n \{ \dimexpr \}

Evaluates the ⟨dimension expression⟩, and leaves the result, expressed in points (pt) in the input stream, with 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:n \{ 1bp \}

leaves 1.00374 in the input stream, i.e. the magnitude of one “big point” when converted to (\TeX) points.
\dim_to_decimal_in_bp:n \{ (dimexpr) \}
\dim_to_decimal_in_sp:n \{ (dimexpr) \}
\dim_to_decimal_in_unit:nn \{ (dimexpr) \} \{ (dimexpr) \}
\dim_to_fp:n \{ (dimexpr) \}
\dim_show:N \{ (dimension) \}
\dim_show:c

25.8 Viewing dim variables
\dim_show:n \{<dimension expression>\}
Displays the result of evaluating the <dimension expression> on the terminal.

\dim_log:N \dim_log:c
\dim_log:n \{<dimension expression>\}
Writes the value of the <dimension> in the log file.

### 25.9 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.

### 25.10 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 \texttt{\LaTeX}\texttt{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_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 \texttt{\LaTeX}\texttt{3}-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

### 25.11 Creating and initialising skip variables

\skip_new:N \skip_new:c
\skip_new:N \{skip\}
Creates a new \texttt{\LaTeX}\texttt{3} \langle\texttt{skip}\rangle or raises an error if the name is already taken. The declaration is global. The \langle\texttt{skip}\rangle is initially equal to 0 pt.
\skip_const:Nn \skip_const:cn
\skip_zero:N \skip_zero:c \skip_gzero:N \skip_gzero:c
\skip_zero_new:N \skip_zero_new:c \skip_gzero_new:N \skip_gzero_new:c
\skip_if_exist_p:N \skip_if_exist_p:c \skip_if_exist:N \skip_if_exist:c
\skip_add:Nn \skip_add:cn \skip_gadd:Nn \skip_gadd:cn
\skip_set:Nn \skip_set:cn \skip_gset:Nn \skip_gset:cn
\skip_sub:Nn \skip_sub:cn \skip_gsub:Nn \skip_gsub:cn
\skip_set_eq:NN \skip_set_eq:(cN|Nc|cc) \skip_gset_eq:NN \skip_gset_eq:(cN|Nc|cc)
\skip_sub:NN \skip_sub:cn \skip_gsub:NN \skip_gsub:cn

\skip_const:Nn \skip (skip) \{\langle \text{skip expression} \rangle \}

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.

\skip_zero:N \skip (skip)
\skip_zero:c
\skip_gzero:N \skip_gzero:c

\skip_zero_new:N \skip (skip)
\skip_zero_new:c \skip_gzero_new:N \skip_gzero_new:c

Ensures that the \langle \text{skip} \rangle exists globally by applying \skip_new:N if necessary, then applies \skip_(g)zero:N to leave the \langle \text{skip} \rangle set to zero.

\skip_if_exist_p:N \skip (skip)
\skip_if_exist_p:c \skip_if_exist:N \skip_if_exist:c

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.

25.12 Setting skip variables

\skip_add:Nn \skip (skip) \{\langle \text{skip expression} \rangle \}
\skip_add:cn \skip_gadd:Nn \skip_gadd:cn

 Adds the result of the \langle \text{skip expression} \rangle to the current content of the \langle \text{skip} \rangle.

\skip_set:Nn \skip (skip) \{\langle \text{skip expression} \rangle \}
\skip_set:cn \skip_gset:Nn \skip_gset:cn

 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).

\skip_set_eq:NN \skip (\text{skip}_1) \langle \text{skip} \rangle \langle \text{skip}_2 \rangle
\skip_set_eq:(cN|Nc|cc) \skip_gset_eq:NN \skip_gset_eq:(cN|Nc|cc)

 Sets the content of \langle \text{skip}_1 \rangle equal to that of \langle \text{skip}_2 \rangle.

\skip_sub:Nn \skip (skip) \{\langle \text{skip expression} \rangle \}
\skip_sub:cn \skip_gsub:Nn \skip_gsub:cn

 Subtracts the result of the \langle \text{skip expression} \rangle from the current content of the \langle \text{skip} \rangle.
25.13 Skip expression conditionals

\skip_if_eq_p:nn
\skip_if_eq:nnTF

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, \textit{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:nTF

Evaluates the \langle skip expression\rangle as described for \skip_eval:n, and then tests if all of its components are finite.

25.14 Using skip expressions and variables

\skip_eval:n
\skip_use:N
\skip_use:c

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 \textit{not} an \langle internal glue\rangle.

\skip_show:N
\skip_show:c
\skip_show:n

Displays the result of evaluating the \langle skip expression\rangle on the terminal.

25.15 Viewing skip variables

\skip_show:N
\skip_show:c

Displays the value of the \langle skip\rangle on the terminal.

\skip_show:n

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

Writes the value of the \langle skip \rangle in the log file.

\skip_log:n \{\langle skip expression \rangle\}

Writes the result of evaluating the \langle skip expression \rangle in the log file.

## 25.16  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.

## 25.17  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.

## 25.18  Inserting skips into the output

\skip_horizontal:N \skip_horizontal:c \skip_horizontal:n

Updated: 2011-10-22

\skip_horizontal: N \langle skip \rangle
\skip_horizontal: n \{\langle skipexpr \rangle\}

Inserts a horizontal \langle skip \rangle into the current list. The argument can also be a \langle 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 \{(skipexpr)\}

Inserts a vertical \langle skip \rangle into the current list. The argument can also be a \langle dim \rangle.

\TeX\ hackers\ note:\ \skip_vertical:N \ is \ the \ \TeX\ primitive \\vskip\ renamed.

\section{Creating and initialising muskip variables}
\muskip_new:N \muskip_new:c
\muskip_const:Nn \muskip_const:cn
\muskip_zero:N \muskip_zero:c \muskip_gzero:N \muskip_gzero:c
\muskip_zero_new:N \muskip_zero_new:c \muskip_gzero_new:N \muskip_gzero_new:c
\muskip_if_exist_p:N \muskip_if_exist_p:c \muskip_if_exist:N \muskip_if_exist:c\ ⋆
\muskip_if_exist:NTF \muskip_if_exist:c\ ⋆
\muskip_if_exist:NTF \muskip_if_exist:CN\ ⋆
\muskip_if_exist:NTF \muskip_if_exist:CN\ ⋆

\muskip_new:N \muskip
\muskip_new:N \muskip

Creates a new \langle muskip \rangle or raises an error if the name is already taken. The declaration
is global. The \langle muskip \rangle is initially equal to 0\,\mu.

\muskip_const:Nn \muskip \{(muskip expression)\}

Creates a new constant \langle muskip \rangle or raises an error if the name is already taken. The
value of the \langle muskip \rangle is set globally to the \langle muskip expression \rangle.

\muskip_zero:N \muskip
\muskip_zero:c \muskip_gzero:N \muskip_gzero:c

Sets \langle muskip \rangle to 0\,\mu.

\muskip_zero_new:N \muskip
\muskip_zero_new:c \muskip_gzero_new:N \muskip_gzero_new:c
\muskip_if_exist_p:N \muskip_if_exist:NTF \muskip_if_exist:CTF
\muskip_if_exist_p:c \muskip_if_exist:CN\ ⋆
\muskip_if_exist:NTF \muskip_if_exist:CN\ ⋆
\muskip_if_exist:NTF \muskip_if_exist:CN\ ⋆
\muskip_if_exist:NTF \muskip_if_exist:CN\ ⋆

\muskip_if_exist_p:N \muskip
\muskip_if_exist_p:c \muskip
\muskip_if_exist:NTF \muskip
\muskip_if_exist:CTF

Tests whether the \langle muskip \rangle is currently defined. This does not check that the \langle muskip \rangle
really is a muskip variable.

\section{Setting muskip variables}
\muskip_add:Nn \muskip_add:cn
\muskip_gadd:Nn \muskip_gadd:cn
\muskip_add:Nn \muskip
\muskip_add:cn
\muskip_gadd:Nn \muskip
\muskip_gadd:cn

Updated: 2011-10-22

\muskip_add:Nn \muskip \{(muskip expression)\}

Adds the result of the \langle muskip expression \rangle to the current content of the \langle muskip \rangle.
Sets \langle \text{muskip} \rangle \text{ to the value of } \langle \text{muskip expression} \rangle, \text{ which must evaluate to a math length with units and may include a rubber component} \text{ (for example } 1 \text{ mu plus } 0.5 \text{ mu).}

\text{Updated: 2011-10-22}

Sets the content of \langle \text{muskip}_1 \rangle \text{ equal to that of } \langle \text{muskip}_2 \rangle.

\text{Updated: 2011-10-22}

Subtracts the result of the \langle \text{muskip expression} \rangle \text{ from the current content of the } \langle \text{muskip} \rangle.

\text{Updated: 2011-10-22}

25.21 \text{ Using muskip expressions and variables}

Evaluates the \langle \text{muskip expression} \rangle, \text{ expanding any skips and token list variables within the } \langle \text{expression} \rangle \text{ to their content (without requiring } \mu \text{kip}_{\text{use:N}}/\tl_{\text{use:N}} \text{) and applying the standard mathematical rules. The result of the calculation is left in the input stream as a } \langle \text{muglue denotation} \rangle \text{ after two expansions. This is expressed in } \mu, \text{ and requires suitable termination if used in a TeX-style assignment as it is not an } \langle \text{internal muglue} \rangle.

\text{TEXhacker note: } \mu \text{kip}_{\text{use:N}} \text{ is the TeX primitive } \texttt{\the}; \text{ this is one of several E\TeX3 names for this primitive.}

25.22 \text{ Viewing muskip variables}

\text{Displays the value of the } \langle \text{muskip} \rangle \text{ on the terminal.}
25.23 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.

25.24 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.

25.25 Primitive conditional

\if_dim:w *
\if_dim:w \{dimen_1\} \{relation\} \{dimen_2\}
\{true code\}
\else:
\{false\}
\fi:

Compare two dimensions. The \{relation\} is one of \(<\), \(=\) or \(>\) with category code 12.

\TeXhackers note: \texHackersNoteThis is the \TeX primitive \ifdim.
Chapter 26

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

```latex
\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:
}\group_begin:
\keys_set:nn { mymodule } { #1 }
% Main code for \MyModuleMacro
\group_end:
```

Key names may contain any tokens, as they are handled internally using `\tl_to_str:n`. As discussed in section 26.2, 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

```latex
\tl_set:Nn \l_mymodule_tmp_tl { key }
\keys_define:nn { mymodule }
  { \l_mymodule_tmp_tl .code:n = code }
```

creates a key called `\l_mymodule_tmp_tl`, and not one called `key`.

### 26.1 Creating keys

\keys_define:nn\keys_define:nn

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

```latex
\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
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 }
{ 
  keyname .value_required:n = true, 
  keyname .code:n = Some-code-using-#1 
}

Note that with the exception of the special \texttt{undefine} property, all key properties define the key within the current \LaTeX{} scope.

\bool_set:N \bool_set:c \bool_gset:N \bool_gset:c
\bool_set_inverse:N \bool_set_inverse:c \bool_gset_inverse:N \bool_gset_inverse:c
\clist_set:N \clist_set:c \clist_gset:N \clist_gset:c

\choice: \choice:

\choices:nn \choices:(Vn|on|xn)

\clist_set:N \clist_set:c \clist_gset:N \clist_gset:c

\texttt{\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 }
{ 
  keyname .value_required:n = true, 
  keyname .code:n = Some-code-using-#1 
}

Note that with the exception of the special \texttt{undefine} property, all key properties define the key within the current \LaTeX{} scope.
Stores the (code) for execution when (key) is used. The (code) can include one parameter (#1), which will be the (value) given for the (key).

Defines (key) to set (control sequence) to have (arg. spec.) and replacement text (value).

Creates a (default) value for (key), which is used if no value is given. This will be used if only the key name is given, but not if a blank (value) is given:

\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.

Defines (key) to set (dimension) to (value) (which must a dimension 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.

Defines (key) to set (floating point) to (value) (which must a floating point 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.

Defines (key) as belonging to the (groups) declared. Groups provide a “secondary axis” for selectively setting keys, and are described in Section 26.6.
\(\text{\texttt{\textbackslash keys\_define:n}}\ \{\ \texttt{foo} \}\ \{\ \texttt{test} \ .\texttt{code:n} = \texttt{\textbackslash tl\_show:n\ \{\#1\}}\}\ \{\ \texttt{bar} \ .\texttt{inherit:n} = \texttt{foo}\}\)

setting

\(\text{\texttt{\textbackslash keys\_set:nn}}\ \{\ \texttt{bar} \}\ \{\ \texttt{test} = a\}\)

will be equivalent to

\(\text{\texttt{\textbackslash keys\_set:nn}}\ \{\ \texttt{foo} \}\ \{\ \texttt{test} = a\}\)

\(\text{\texttt{\textbackslash keys\_define:n}}\ \{\ \texttt{foo} \}\ \{\ \texttt{test} \ .\texttt{code:n} = \texttt{\textbackslash tl\_show:n\ \{\#1\}}\}\ \{\ \texttt{bar} \ .\texttt{inherit:n} = \texttt{foo} \}\)

\(\text{\texttt{\textbackslash keys\_set:nn}}\ \{\ \texttt{bar} \}\ \{\ \texttt{test} = a\}\)

will be equivalent to

\(\text{\texttt{\textbackslash keys\_set:nn}}\ \{\ \texttt{foo} \}\ \{\ \texttt{test} = a\}\)

\(\text{\texttt{\textbackslash keys\_define:n}}\ \{\ \texttt{foo} \}\ \{\ \texttt{test} \ .\texttt{code:n} = \texttt{\textbackslash tl\_show:n\ \{\#1\}}\}\ \{\ \texttt{bar} \ .\texttt{inherit:n} = \texttt{foo} \}\)

\(\text{\texttt{\textbackslash keys\_set:nn}}\ \{\ \texttt{bar} \}\ \{\ \texttt{test} = a\}\)

will be equivalent to

\(\text{\texttt{\textbackslash keys\_set:nn}}\ \{\ \texttt{foo} \}\ \{\ \texttt{test} = a\}\)
26.2 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:
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.

### 26.3 Choice and multiple choice keys

The l3keys 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 `.choice:` property:

```latex
\keys_define:nn { mymodule }
  { key .choice: }
```

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 `.choices:nn` property.

```latex
\keys_define:nn { mymodule }
  { key .choices: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.
    }
  }
```

The index `\l_keys_choice_int` in the list of choices starts at 1.

Inside the code block for a choice generated using `.choices:nn`, the variables `\l_keys_choice_tl` and `\l_keys_choice_int` are available to indicate the name of the current choice, and its position in the comma list. The position is indexed from 1. Note that, as with standard key code generated using `.code:n`, the value passed to the key (i.e. the choice name) is also available as `#1`.
On the other hand, it is sometimes useful to create choices which use entirely different code from one another. This can be achieved by setting the `.choice:` property of a key, then manually defining sub-keys.

\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,
}

It is possible to mix the two methods, but manually-created choices should not use \l_keys_choice_tl or \l_keys_choice_int. These variables do not have defined behaviour when used outside of code created using .choices:nn (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 `unknown` choice. The general behavior of the `unknown` key is described in Section 26.5. A typical example in the case of a choice would be to issue a custom error message:

\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:nnxxx { mymodule } { unknown-choice }
    { key } % Name of choice key
    { choice-a , choice-b , choice-c } % Valid choices
    { \exp_not:n {#1} } % Invalid choice given
%
%
}

Multiple choices are created in a very similar manner to mutually-exclusive choices, using the properties `.multchoicen` and `.multchoicess:nn`. As with mutually exclusive choices, multiple choices are defined as sub-keys. Thus both

\keys_define:nn { mymodule }
{
    key .multchoicen =
    { 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.
    }
}
and

\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,
}

are valid.

When a multiple choice key is set

\keys_set:nn { mymodule }
{
  key = { a , b , c } % 'key' defined as a multiple choice
}

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.

\section{Setting keys}

\keys_set:nn \keys_set:(nV|nv|no)
Updated: 2017-11-14

\keys_set:nn \keys_set:\langle\text{module}\rangle\{\langle\text{keyval list}\rangle\}

Parses the \langle\text{keyval list}\rangle, and sets those keys which are defined for \langle\text{module}\rangle. The behaviour on finding an unknown key can be set by defining a special \texttt{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.

The value is 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.

The path of the key is a “full” description of the key, and is unique for each key. It
consists of the module and full key name, thus for example

\keys_set:nn { mymodule } { key-a = some-value }

has path mymodule/key-a while

\keys_set:nn { mymodule } { subset / key-a = some-value }

has path mymodule/subset/key-a. This information is stored in \l_keys_path_str.

The name of the key is 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.

### 26.5 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 \module, and simply ignore other
keys. The \keys_set_known:nn function parses the \keyval list, and sets those keys
which are defined for \module. Any keys which are unknown are not processed further
by the parser. In addition, \keys_set_known:nnN stores the key–value pairs in the (tl)
in comma-separated form (i.e. an edited version of the \keyval list). When a \root is
given (\keys_set_known:nnN), 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.
26.6 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 \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 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.

231
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.

26.7 Utility functions for keys

Tests if the ⟨key⟩ exists for ⟨module⟩, i.e. if any code has been defined for ⟨key⟩.

Tests if the ⟨choice⟩ is defined for the ⟨key⟩ within the ⟨module⟩, i.e. if any code has been defined for ⟨key⟩/⟨choice⟩. The test is false if the ⟨key⟩ itself is not defined.

Displays in the terminal the information associated to the ⟨key⟩ for a ⟨module⟩, including the function which is used to actually implement it.

Writes in the log file the information associated to the ⟨key⟩ for a ⟨module⟩. See also \keys_show:nn which displays the result in the terminal.

26.8 Low-level interface for parsing key–val lists

To re-cap from earlier, a key–value list is input of the form

KeyOne = ValueOne ,
KeyTwo = ValueTwo ,
KeyThree

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 ⟨key–value list⟩ into ⟨keys⟩ and associated ⟨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 # tokens or expand any input. Active tokens = and , 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

\[
\text{key} = \{\text{value here}\},
\]

and

\[
\text{key} = \text{value here},
\]

are treated identically.

\keyval_parse:nnn \keyval_parse:nnn \{(code\textsubscript{1})\} \{(code\textsubscript{2})\} \{\text{key–value list}\}

Parses the \{key–value list\} into a series of \{keys\} and associated \{values\}, or keys alone (if no \{value\} was given). \{code\textsubscript{1}\} receives each \{key\} (with no \{value\}) as a trailing brace group, whereas \{code\textsubscript{2}\} is appended by two brace groups, the \{key\} and \{value\}. The order of the \{keys\} in the \{key–value list\} is preserved. Thus

\keyval_parse:nnn

{ \use_none:nn { code 1 } }
{ \use_none:nn { code 2 } }
{ key1 = value1, key2 = value2, key3 = , key4 }

is converted into an input stream

\use_none:nn { code 2 } { key1 } { value1 }
\use_none:nn { code 2 } { key2 } { value2 }
\use_none:nn { code 2 } { key3 } {}
\use_none:nn { code 1 } { 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. If you need exactly the output shown above, you’ll need to either \texttt{x-type} or \texttt{e-type} expand the function.

\textbf{\LaTeX hack note:} The result of each list element is returned within \texttt{\exp_not:n}, which means that the converted input stream does not expand further when appearing in an \texttt{x-type} or \texttt{e-type} argument expansion.
\keyval_parse:NNn \keyval_parse:NNn (function_1) (function_2) \{\langle key–value list\rangle\}

Parses the \langle key–value list\rangle into a series of \langle keys\rangle and associated \langle values\rangle, or keys alone (if no \langle value\rangle was given). \langle function_1\rangle should take one argument, while \langle function_2\rangle should absorb two arguments. After \keyval_parse:NNn has parsed the \langle key–value list\rangle, \langle function_1\rangle is used to process keys given with no value and \langle function_2\rangle is used to process keys given with a value. The order of the \langle keys\rangle in the \langle key–value list\rangle 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 \langle key\rangle and \langle value\rangle, then one outer set of braces is removed from the \langle key\rangle and \langle value\rangle as part of the processing.

This shares the implementation of \keyval_parse:nnn, the difference is only semantically.

\TeXhackers 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.
Chapter 27

The \texttt{l3intarray} package: fast global integer arrays

27.1 \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 $2^{30} - 1$ (i.e. one power lower than the usual \texttt{c_max_int} 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
\ \langle intarray var \rangle \{\langle size \rangle\}
\end{verbatim}

Evaluates the integer expression \texttt{\langle size \rangle} and allocates an \texttt{\langle integer array variable \rangle} 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 \intarray_count:c
\ \langle intarray var \rangle
\end{verbatim}

Expands to the number of entries in the \texttt{\langle integer array variable \rangle}. Contrarily to \texttt{\seq_\texttt{-count:N}} this is performed in constant time.

\begin{verbatim}
\intarray_gset:Nnn \intarray_gset:cn \intarray_gset:cn
\ \langle intarray var \rangle \{\langle position \rangle\} \{\langle value \rangle\}
\end{verbatim}

Stores the result of evaluating the integer expression \texttt{\langle value \rangle} into the \texttt{\langle integer array variable \rangle} at the (integer expression) \texttt{\langle position \rangle}. If the \texttt{\langle position \rangle} is not between 1 and the \texttt{\intarray_count:N}, or the \texttt{\langle value \rangle}’s absolute value is bigger than $2^{30} - 1$, an error occurs. Assignments are always global.
\intarray_const_from_clist:Nn \intarray_const_from_clist:Nn \intarray_var \intexpr_clist

\intarray_const_from_clist:cn

New: 2018-05-04

Creates a new constant (integer array variable) or raises an error if the name is already taken. The (integer array variable) is set (globally) to contain as its items the results of evaluating each (integer expression) in the (comma list).

\intarray_gzero:N \intarray_gzero:cn

New: 2018-05-04

\intarray_gzero:N \intarray_var

Sets all entries of the (integer array variable) to zero. Assignments are always global.

\intarray_item:Nn \intarray_item:cn *

New: 2018-03-29

\intarray_item:Nn \intarray_var \{ \text{position}\}

Expands to the integer entry stored at the (integer expression) (position) in the (integer array variable). If the (position) is not between 1 and the \intarray_count:N, an error occurs.

\intarray_rand_item:N \intarray_rand_item:cn *

New: 2018-05-05

\intarray_rand_item:N \intarray_var

Selects a pseudo-random item of the (integer array). If the (integer array) is empty, produce an error.

\intarray_show:N \intarray_show:cn \intarray_log:N \intarray_log:cn

New: 2018-05-04

\intarray_show:N \intarray_var \intarray_log:N \intarray_var

Displays the items in the (integer array variable) in the terminal or writes them in the log file.

27.1.1 Implementation notes

It is a wrapper around the \fontdimen primitive, used to store arrays of integers (with a restricted range: absolute value at most $2^{30} - 1$). In contrast to \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 \fontdimen stores dimensions but the \l3intarray package transparently converts these from/to integers. Assignments are always global.

While LuaTeX’s memory is extensible, other engines can “only” deal with a bit less than $4 \times 10^6$ entries in all \fontdimen arrays combined (with default \TeX Live settings).
Chapter 28

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 $\neg x$, 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: $\arcsin(x)$, $\arccos(x)$, $\arctan(x)$, $\arccot(x)$, $\arcsec(x)$, $\arccsc(x)$ giving a result in radians, and $\arcsind(x)$, $\arccosd(x)$, $\arctand(x)$, $\arccotd(x)$, $\arcsecd(x)$, $\arccscd(x)$ giving a result in degrees.

(not yet) Hyperbolic functions and their inverse functions: $\sinh(x)$, $\cosh(x)$, $\tanh(x)$, $\coth(x)$, $\text{sech}(x)$, $\text{csch}(x)$, and $\text{arsinh}(x)$, $\text{arccosh}(x)$, $\text{artanh}(x)$, $\text{arcoth}(x)$, $\text{arsech}(x)$, $\text{arcsch}(x)$.

- 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., pc is 12.
• Automatic conversion (no need for $\langle \text{type}\rangle_{\text{use}:N}$) 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 28.9.1 for a description of what a floating point is, section 28.9.2 for details about how an expression is parsed, and section 28.9.3 to know what the various operations do. Some operations may raise exceptions (error messages), described in section 28.7.

An example of use could be the following.

\LaTeX{} can now compute: $\frac{\sin (3.5)}{2} + 2 \cdot 10^{-3}$ \[ \text{\LaTeX{} can now compute: $\frac{\sin (3.5)}{2} + 2 \cdot 10^{-3}$} \]

The operation $\text{round}$ 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 ( 2.3 ^ 5 ) }\end{document}

See the documentation of siunitx for various options of \num.
28.1 Creating and initialising floating point variables

\texttt{\fp_new:N (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.

\texttt{\fp_new:N (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)}.

\texttt{\fp_zero:N (fp var)}

Sets the \texttt{(fp var)} to +0.

\texttt{\fp_zero:N (fp var)}

Ensures that the \texttt{(fp var)} exists globally by applying \texttt{\fp_new:N} if necessary, then applies \texttt{\fp\_\(g\)zero:N} to leave the \texttt{(fp var)} set to +0.

28.2 Setting floating point variables

\texttt{\fp_set:Nn (fp var) \{\texttt{floating point expression}\}}

Sets \texttt{(fp var)} equal to the result of computing the \texttt{(floating point expression)}.

\texttt{\fp_set_eq:NN \texttt{\texttt{fp var}_1} \texttt{\texttt{fp var}_2}}

Sets the floating point variable \texttt{(fp var\texttt{_1})} equal to the current value of \texttt{(fp var\texttt{\_2})}.

\texttt{\fp_add:Nn (fp var) \{\texttt{floating point expression}\}}

Adds the result of computing the \texttt{(floating point expression)} to the \texttt{(fp var)}. This also applies if \texttt{(fp var)} and \texttt{(floating point expression)} evaluate to tuples of the same size.
\texttt{\textbackslash fp\_sub:Nn} \texttt{\textbackslash fp\_sub:cn} 
\texttt{\textbackslash fp\_gsub:Nn} \texttt{\textbackslash fp\_gsub:cn} Updated: 2012-06-08

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.

\section{Using floating points}

\texttt{\textbackslash fp\_eval:n} \texttt{\textbackslash fp\_eval:cn} Updated: 2012-05-08

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 $\pm \infty$ and NaN trigger an “invalid operation” exception. For a tuple, each item is converted using \texttt{\textbackslash fp\_eval:n} and they are combined as $\langle fp_1 \rangle, \langle fp_2 \rangle, \ldots, \langle fp_n \rangle$ if $n > 1$ and $\langle fp_1 \rangle$, or {} for fewer items. This function is identical to \texttt{\textbackslash fp\_to\_decimal:n}.

\texttt{\textbackslash fp\_sign:n} \texttt{\textbackslash fp\_sign:cn} Updated: 2012-05-08 Updated: 2012-07-08

Evaluates the \textit{(fpexpr)} and leaves its sign in the input stream using \texttt{\textbackslash fp\_eval:n} \{\texttt{\textbackslash sign(\langle result\rangle)}}: +1 for positive numbers and for $+\infty$, $-1$ for negative numbers and for $-\infty$, $\pm 0$ for $\pm 0$. If the operand is a tuple or is NaN, then “invalid operation” occurs and the result is 0.

\texttt{\textbackslash fp\_to\_decimal:N} \texttt{\textbackslash fp\_to\_decimal:c} \texttt{\textbackslash fp\_to\_decimal:n} \texttt{\textbackslash updated\_2012-05-08}

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 $\pm \infty$ and NaN trigger an “invalid operation” exception. For a tuple, each item is converted using \texttt{\textbackslash fp\_to\_decimal:n} and they are combined as $\langle fp_1 \rangle, \langle fp_2 \rangle, \ldots, \langle fp_n \rangle$ if $n > 1$ and $\langle fp_1 \rangle$, or {} for fewer items.

\texttt{\textbackslash fp\_to\_dim:N} \texttt{\textbackslash fp\_to\_dim:c} \texttt{\textbackslash fp\_to\_dim:n} \texttt{\textbackslash updated\_2016-03-22}

Evaluates the \textit{(floating point expression)} and expresses the result as a dimension (in pt) suitable for use in dimension expressions. The output is identical to \texttt{\textbackslash fp\_to\_decimal:n} with an additional trailing 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 NaN, trigger an “invalid operation” exception.

\texttt{\textbackslash fp\_to\_int:N} \texttt{\textbackslash fp\_to\_int:c} \texttt{\textbackslash fp\_to\_int:n} \texttt{\textbackslash updated\_2012-07-08}

Evaluates the \textit{(floating point expression)}, 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 NaN, trigger an “invalid operation” exception.
\fp_to_scientific:N \fp_to_scientific:c \fp_to_scientific:n

Evaluates the \textit{floating point expression} and expresses the result in scientific notation:

\[ (\text{optional } -) (\text{digit}) . (15 \text{ digits}) \text{e} (\text{optional sign}) (\text{exponent}) \]

The leading \textit{digit} is non-zero except in the case of \pm 0. The values \pm \infty and NaN trigger an “invalid operation” exception. Normal category codes apply: thus the \textit{e} is category code 11 (a letter). For a tuple, each item is converted using \fp_to_scientific:n and they are combined as \((\langle fp_1 \rangle, \langle fp_2 \rangle, \ldots \langle fp_n \rangle)\) if \(n > 1\) and \((\langle fp_1 \rangle,)\) or () for fewer items.

\fp_to_tl:N \fp_to_tl:c \fp_to_tl:n

Evaluates the \textit{floating point expression} and expresses the result in (almost) the shortest possible form. Numbers in the ranges \((0, 10^{-3})\) and \([10^{16}, \infty)\) are expressed in scientific notation with trailing zeros trimmed and no decimal separator when there is a single significant digit (this differs from \fp_to_scientific:n). Numbers in the range \([10^{-3}, 10^{16}]\) are expressed in a decimal notation without exponent, with trailing zeros trimmed, and no decimal separator for integer values (see \fp_to_decimal:n. Negative numbers start with \textit{-}. The special values \pm 0, \pm \infty and NaN are rendered as 0, -0, \textit{inf}, \textit{-inf}, and \textit{nan} respectively. Normal category codes apply and thus \textit{inf} or \textit{nan}, if produced, are made up of letters. For a tuple, each item is converted using \fp_to_tl:n and they are combined as \((\langle fp_1 \rangle, \langle fp_2 \rangle, \ldots \langle fp_n \rangle)\) if \(n > 1\) and \((\langle fp_1 \rangle,)\) or () for fewer items. This function is identical to \fp_to_decimal:n.

\fp_use:N \fp_use:c

Inserts the value of the \textit{fp var} into the input stream 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. Integers are expressed without a decimal separator. The values \pm \infty and NaN trigger an “invalid operation” exception. For a tuple, each item is converted using \fp_to_decimal:n and they are combined as \((\langle fp_1 \rangle, \langle fp_2 \rangle, \ldots \langle fp_n \rangle)\) if \(n > 1\) and \((\langle fp_1 \rangle,)\) or () for fewer items. This function is identical to \fp_to_decimal:n.

\fp_if_exist_p:N \fp_if_exist_p:c \fp_if_exist:NTF \fp_if_exist:cTF

Tests whether the \textit{fp var} is currently defined. This does not check that the \textit{fp var} really is a floating point variable.
\fp_compare_p:nNn * \fp_compare_p:nNn \{\textit{fpexpr}_1\} \{\textit{relation}\} \{\textit{fpexpr}_2\}
\fp_compare:nNnTF \{\textit{fpexpr}_1\} \{\textit{relation}\} \{\textit{fpexpr}_2\} \{\textit{true code}\} \{\textit{false code}\}

Compares the \textit{fpexpr}_1 and the \textit{fpexpr}_2, and returns \texttt{true} if the \textit{relation} is obeyed. Two floating points \textit{x} and \textit{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 NaN or is a tuple, unless they are equal tuples. Note that a NaN is distinct from any value, even another NaN, hence \texttt{x = x} is not true for a NaN. To test if a value is NaN, compare it to an arbitrary number with the “not ordered” relation.

\begin{verbatim}
\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 NaN). At present any other comparison with tuples yields \texttt{?} (not ordered). This is experimental.

This function is less flexible than \fp_compare:nTF but slightly faster. It is provided for consistency with \int_compare:nNnTF and \dim_compare:nNnTF.
\fp_compare_p:n \fp_compare:nTF
\fp_compare_p:n
{
(fexpr_1) (relation_1)
...
(fexpr_N) (relation_N)
(fexpr_{N+1})
}
\fp_compare:nTF
{
(fexpr_1) (relation_1)
...
(fexpr_N) (relation_N)
(fexpr_{N+1})
}
{(true code)} {{false code}}

Evaluates the \textit{(floating point expressions)} as described for \texttt{\fp_eval:n} and compares consecutive result using the corresponding \texttt{(relation)}, namely it compares \texttt{⟨fpexpr_1⟩} and \texttt{⟨fpexpr_2⟩} using the \texttt{(relation_1)}, then \texttt{⟨fpexpr_2⟩} and \texttt{⟨fpexpr_3⟩} using the \texttt{(relation_2)}, until finally comparing \texttt{⟨fpexpr_N⟩} and \texttt{⟨fpexpr_{N+1}⟩} using the \texttt{(relation_N)}. The test yields \texttt{true} if all comparisons are \texttt{true}. Each \texttt{(floating point expression)} is evaluated only once. Contrarily to \texttt{\int_compare:nTF}, all \texttt{(floating point expressions)} are computed, even if one comparison is \texttt{false}. 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. Each \texttt{(relation)} can be any (non-empty) combination of \texttt{<, =, >, and ?}, plus an optional leading \texttt{!} (which negates the \texttt{(relation)}), with the restriction that the \texttt{(relation)} may not start with \texttt{?}, as this symbol has a different meaning (in combination with \texttt{:}) within floating point expressions. The comparison \texttt{x (relation) y} is then \texttt{true} if the \texttt{(relation)} does not start with \texttt{!} and the actual relation \texttt{(<, =, >, or ?)} between \texttt{x} and \texttt{y} appears within the \texttt{(relation)}, or on the contrary if the \texttt{(relation)} starts with \texttt{!} and the relation between \texttt{x} and \texttt{y} does not appear within the \texttt{(relation)}. Common choices of \texttt{(relation)} include \texttt{>=} (greater or equal), \texttt{!=} (not equal), \texttt{!?} or \texttt{<=} (comparable).

This function is more flexible than \texttt{\fp_compare:nNnTF} and only slightly slower.

### 28.5 Floating point expression loops

\texttt{\fp_do_until:nNnn} \texttt{\fp_do_until:nNnn}
\\texttt{(fexpr_1)} \texttt{(relation)} \texttt{(fexpr_2)} \texttt{(code)}

Places the \texttt{(code)} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \texttt{(floating point expressions)} as described for \texttt{\fp_compare:nNnTF}. If the test is \texttt{false} then the \texttt{(code)} is inserted into the input stream again and a loop occurs until the \texttt{(relation)} is \texttt{true}.

\texttt{\fp_do_while:nNnn} \texttt{\fp_do_while:nNnn}
\\texttt{(fexpr_1)} \texttt{(relation)} \texttt{(fexpr_2)} \texttt{(code)}

Places the \texttt{(code)} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \texttt{(floating point expressions)} as described for \texttt{\fp_compare:nNnTF}. If the test is \texttt{true} then the \texttt{(code)} is inserted into the input stream again and a loop occurs until the \texttt{(relation)} is \texttt{false}.
\fp_until_do:nNnn \{\text{relation}\} \{\text{code}\}

Evaluates the relationship between the two \emph{floating point expressions} as described for \texttt{fp_compare:nNnTF}, and then places the \texttt{code} in the input stream if the \texttt{relation} is \texttt{false}. After the \texttt{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \texttt{true}.

\fp_while_do:nNnn \{\text{relation}\} \{\text{code}\}

Evaluates the relationship between the two \emph{floating point expressions} as described for \texttt{fp_compare:nNnTF}, and then places the \texttt{code} in the input stream if the \texttt{relation} is \texttt{true}. After the \texttt{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \texttt{false}.

\fp_do_until:nn \{\text{relation}\} \{\text{code}\}

Places the \texttt{code} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \emph{floating point expressions} as described for \texttt{fp_compare:nNnTF}. If the test is \texttt{false} then the \texttt{code} is inserted into the input stream again and a loop occurs until the \texttt{relation} is \texttt{true}.

\fp_do_while:nn \{\text{relation}\} \{\text{code}\}

Places the \texttt{code} in the input stream for \TeX{} to process, and then evaluates the relationship between the two \emph{floating point expressions} as described for \texttt{fp_compare:nNnTF}. If the test is \texttt{true} then the \texttt{code} is inserted into the input stream again and a loop occurs until the \texttt{relation} is \texttt{false}.

\fp_until_do:nn \{\text{relation}\} \{\text{code}\}

Evaluates the relationship between the two \emph{floating point expressions} as described for \texttt{fp_compare:nNnTF}, and then places the \texttt{code} in the input stream if the \texttt{relation} is \texttt{false}. After the \texttt{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \texttt{true}.

\fp_while_do:nn \{\text{relation}\} \{\text{code}\}

Evaluates the relationship between the two \emph{floating point expressions} as described for \texttt{fp_compare:nNnTF}, and then places the \texttt{code} in the input stream if the \texttt{relation} is \texttt{true}. After the \texttt{code} has been processed by \TeX{} the test is repeated, and a loop occurs until the test is \texttt{false}.
\fp_step_function:nnnN \fp_step_function:nnnc

New: 2016-11-21
Updated: 2016-12-06

This function first evaluates the \langle initial value \rangle, \langle step \rangle and \langle final value \rangle, each of which should be a floating point expression evaluating to a floating point number, not a tuple. The \langle function \rangle is then placed in front of each \langle value \rangle from the \langle initial value \rangle to the \langle final value \rangle in turn (using \langle step \rangle between each \langle value \rangle). The \langle step \rangle must be non-zero. If the \langle step \rangle is positive, the loop stops when the \langle value \rangle becomes larger than the \langle final value \rangle. If the \langle step \rangle is negative, the loop stops when the \langle value \rangle becomes smaller than the \langle final value \rangle. The \langle function \rangle should absorb one numerical argument. For example

\cs_set:Npn \my_func:n #1 { \[I saw #1\] \quad }
\fp_step_function:nnn { 1.0 } { 0.1 } { 1.5 } \my_func:n

would print

[I saw 1.0] [I saw 1.1] [I saw 1.2] [I saw 1.3] [I saw 1.4] [I saw 1.5]

\TeX\ hackers note: Due to rounding, it may happen that adding the \langle step \rangle to the \langle value \rangle does not change the \langle value \rangle; such cases give an error, as they would otherwise lead to an infinite loop.

\fp_step_inline:nnnn

New: 2016-11-21
Updated: 2016-12-06

This function first evaluates the \langle initial value \rangle, \langle step \rangle and \langle final value \rangle, all of which should be floating point expressions evaluating to a floating point number, not a tuple. Then for each \langle value \rangle from the \langle initial value \rangle to the \langle final value \rangle in turn (using \langle step \rangle between each \langle value \rangle), the \langle code \rangle is inserted into the input stream with \#1 replaced by the current \langle value \rangle. Thus the \langle code \rangle should define a function of one argument (\#1).

\fp_step_variable:nnnNn

New: 2017-04-12

This function first evaluates the \langle initial value \rangle, \langle step \rangle and \langle final value \rangle, all of which should be floating point expressions evaluating to a floating point number, not a tuple. Then for each \langle value \rangle from the \langle initial value \rangle to the \langle final value \rangle in turn (using \langle step \rangle between each \langle value \rangle), the \langle code \rangle is inserted into the input stream, with the \langle tl var \rangle defined as the current \langle value \rangle. Thus the \langle code \rangle should make use of the \langle tl var \rangle.

28.6 Some useful constants, and scratch variables

\c_zero_fp
\c_minus_zero_fp

New: 2012-05-08

Zero, with either sign.

\c_one_fp

New: 2012-05-08

One as an \texttt{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 $1^\circ$ 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 \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 \LaTeX3-defined function. However, they may be overwritten by other non-kernel code and so should only be used for short-term storage.

28.7 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 $10 ^ {1e9999}$. 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 $\pm 0$.

- **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., \texttt{\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 $\pm \infty$.  

246
(not yet) 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\.

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.

\begin{verbatim}
\fp\_trap:nn \{\langle\text{exception}\rangle\} \{\langle\text{trap type}\rangle\}
\end{verbatim}

All occurrences of the \texttt{\langle\text{exception}\rangle} (overflow, underflow, invalid\_operation or division\_by\_zero) within the current group are treated as \texttt{\langle\text{trap type}\rangle}, which can be:

- none: the \texttt{\langle\text{exception}\rangle} will be entirely ignored, and leave no trace;
- flag: the \texttt{\langle\text{exception}\rangle} will turn the corresponding flag on when it occurs;
- error: additionally, the \texttt{\langle\text{exception}\rangle} 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.

\begin{verbatim}
flag\_fp\_overflow
flag\_fp\_underflow
flag\_fp\_invalid\_operation
flag\_fp\_division\_by\_zero
\end{verbatim}

\section{28.8 Viewing floating points}

\begin{verbatim}
\fp\_show:N \langle\text{fp var}\rangle
\fp\_show:C
\fp\_show:n \{\langle\text{floating point expression}\rangle\}
\end{verbatim}

Evaluates the \texttt{\langle\text{floating point expression}\rangle} and displays the result in the terminal.

\begin{verbatim}
\fp\_log:N \langle\text{fp var}\rangle
\fp\_log:C
\fp\_log:n \{\langle\text{floating point expression}\rangle\}
\end{verbatim}

Evaluates the \texttt{\langle\text{floating point expression}\rangle} and writes the result in the log file.
28.9 Floating point expressions

28.9.1 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 $-10000 \leq n \leq 10000$;
- $\pm0$, 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 or 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 $\pm0$).

The result is thus $\pm0$ 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 (which is faster).

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.
28.9.2 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 (\(\sin, \ln, \text{etc}\)).
- Binary \(^*\) and \(^-\) (right associative).
- Unary +, -, !.
- Implicit multiplication by juxtaposition (\(2\pi\)) when neither factor is in parentheses.
- Binary \(*\) and \(/\), implicit multiplication by juxtaposition with parentheses (for instance \(3(4+5)\)).
- Binary + and -.
- Comparisons \(\geq, \neq, <, \text{etc}\).
- Logical and, denoted by \&\&.
- Logical or, denoted by ||.
- Ternary operator \(?:\) (right associative).
- Comma (to build tuples).

The precedence of operations can be overridden using parentheses. In particular, the precedence of juxtaposition implies that

\[\frac{1}{2\pi} = \frac{1}{(2\pi)},\]
\[\frac{1}{2\pi}(\pi + \pi) = (2\pi)^{-1}(\pi + \pi) \simeq 1,\]
\[\sin 2\pi = \sin(2\pi) \pi \neq 0,\]
\[2^{\text{max}}(3, 5) = 2^{\text{max}}(3, 5) = 20,\]
\[\text{in}/\text{cm} = (\text{in})/(\text{cm}) = 2.54.\]

Functions are called on the value of their argument, contrarily to \TeX{} macros.

28.9.3 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 \(\pm 0\), 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 (\(?:\), \[\], \&\&, !), by comparisons (\(!\leq\?\)), and by +, -, *, /. 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 \( \langle \text{operand}_2 \rangle \) if \( \langle \text{operand}_1 \rangle \) is true (not ±0), and \( \langle \text{operand}_3 \rangle \) if \( \langle \text{operand}_1 \rangle \) is false (±0). All three \( \langle \text{operands} \rangle \) 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.

\[
\fp_eval:n \{ \langle \text{operand}_1 \rangle \ || \langle \text{operand}_2 \rangle \}
\]

If \( \langle \text{operand}_1 \rangle \) is true (not ±0), use that value, otherwise the value of \( \langle \text{operand}_2 \rangle \). Both \( \langle \text{operands} \rangle \) are evaluated in all cases; they may be tuples. In \( \langle \text{operand}_1 \rangle \ || \langle \text{operand}_2 \rangle \ || \ldots \ || \langle \text{operands}_n \rangle \), the first true (nonzero) \( \langle \text{operand} \rangle \) is used and if all are zero the last one (±0) is used.

\[
\fp_eval:n \{ \langle \text{operand}_1 \rangle \ && \langle \text{operand}_2 \rangle \}
\]

If \( \langle \text{operand}_1 \rangle \) is false (equal to ±0), use that value, otherwise the value of \( \langle \text{operand}_2 \rangle \). Both \( \langle \text{operands} \rangle \) are evaluated in all cases; they may be tuples. In \( \langle \text{operand}_1 \rangle \ && \langle \text{operand}_2 \rangle \ && \ldots \ && \langle \text{operands}_n \rangle \), the first false (±0) \( \langle \text{operand} \rangle \) is used and if none is zero the last one is used.

Each \( \langle \text{relation} \rangle \) consists of a non-empty string of \( <, =, >, \) and \?, optionally preceded by \!, and may not start with \?. This evaluates to +1 if all comparisons \( \langle \text{operand}_i \rangle \ \langle \text{relation}_i \rangle \ \langle \text{operand}_{i+1} \rangle \) are true, and +0 otherwise. All \( \langle \text{operands} \rangle \) are evaluated (once) in all cases. See \texttt{\textbackslash fp_compare:nTF} for details.

\[
\fp_eval:n \{ \langle \text{operand}_1 \rangle + \langle \text{operand}_2 \rangle \}
\]

\[
\fp_eval:n \{ \langle \text{operand}_1 \rangle - \langle \text{operand}_2 \rangle \}
\]

Computes the sum or the difference of its two \( \langle \text{operands} \rangle \). The “invalid operation” exception occurs for \( \infty - \infty \). “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.
\[ \text{\textit{fp_eval:n \{} \langle operand_1 \rangle \star \langle operand_2 \rangle \}} \]
\[ \text{\textit{fp_eval:n \{} \langle operand_1 \rangle \div \langle operand_2 \rangle \}} \]

Computes the product or the ratio of its two \langle operands \rangle. The “invalid operation” exception occurs for \( \infty/\infty \), 0/0, or 0 \* \( \infty \). “Division by zero” occurs when dividing a finite non-zero number by ±0. “Underflow” and “overflow” occur when appropriate. When \langle operand_1 \rangle is a tuple and \langle operand_2 \rangle is a floating point number, each item of \langle operand_1 \rangle is multiplied or divided by \langle operand_2 \rangle. Multiplication also supports the case where \langle operand_1 \rangle is a floating point number and \langle operand_2 \rangle a tuple. Other combinations yield an “invalid operation” exception and a NaN result.

\[ + \text{\textit{fp_eval:n \{} + \langle operand \rangle \}} \]
\[ - \text{\textit{fp_eval:n \{} - \langle operand \rangle \}} \]
\[ ! \text{\textit{fp_eval:n \{} ! \langle operand \rangle \}} \]

The unary + does nothing, the unary - changes the sign of the \langle operand \rangle (for a tuple, of all its components), and ! \langle operand \rangle evaluates to 1 if \langle operand \rangle is false (is ±0) and 0 otherwise (this is the \texttt{not} boolean function). Those operations never raise exceptions.

\[ ** \text{\textit{fp_eval:n \{} \langle operand_1 \rangle ** \langle operand_2 \rangle \}} \]
\[ ^{\star} \text{\textit{fp_eval:n \{} \langle operand_1 \rangle ^{\star} \langle operand_2 \rangle \}} \]

Raises \langle operand_1 \rangle to the power \langle operand_2 \rangle. This operation is right associative, hence 2 ** 2 ** 3 equals 2\(^2 \times 3 = 256\). If \langle operand_1 \rangle is negative or \(-0\) then: the result’s sign is + if the \langle operand_2 \rangle is infinite and \((-1)^p\) if the \langle operand_2 \rangle is \(p/q\) with \(p, q\) integers; the result is \(+0\) if abs(\langle operand_1 \rangle)\(^{\star} \langle operand_2 \rangle\) evaluates to zero; in other cases the “invalid operation” exception occurs because the sign cannot be determined. “Division by zero” occurs when raising ±0 to a finite strictly negative power. “Underflow” and “overflow” occur when appropriate. If either operand is a tuple, “invalid operation” occurs.

\[ \text{\textit{abs \{} \text{abs( \langle fpexpr \rangle ) \}} \]

Computes the absolute value of the \langle fpexpr \rangle. If the operand is a tuple, “invalid operation” occurs. This operation does not raise exceptions in other cases. See also \textit{fp_abs:n}.

\[ \text{\textit{exp \{} \text{exp( \langle fpexpr \rangle ) \}} \]

Computes the exponential of the \langle fpexpr \rangle. “Underflow” and “overflow” occur when appropriate. If the operand is a tuple, “invalid operation” occurs.

\[ \text{\textit{fact \{} \text{fact( \langle fpexpr \rangle ) \}} \]

Computes the factorial of the \langle fpexpr \rangle. If the \langle fpexpr \rangle is an integer between \(-0\) and 3248 included, the result is finite and correctly rounded. Larger positive integers give +\( \infty \) with “overflow”, while fact(+\( \infty \)) = +\( \infty \) and fact(nan) = nan with no exception. All other inputs give NaN with the “invalid operation” exception.

\[ \text{\textit{ln \{} \text{ln( \langle fpexpr \rangle ) \}} \]

Computes the natural logarithm of the \langle fpexpr \rangle. 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.
\texttt{logb} \ast \ \texttt{fp_eval:n \{ logb( \langle fpexpr \rangle ) \}}

Determines the exponent of the \langle \text{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}.

\texttt{max} \ \texttt{fp_eval:n \{ max( \langle fpexpr_1 \rangle, \langle fpexpr_2 \rangle, \ldots ) \}}
\texttt{min} \ \texttt{fp_eval:n \{ min( \langle fpexpr_1 \rangle, \langle fpexpr_2 \rangle, \ldots ) \}}

Evaluates each \langle \text{fpexpr} \rangle and computes the largest (smallest) of those. If any of the \langle \text{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.

\texttt{round} \ \texttt{fp_eval:n \{ round( \langle fpexpr \rangle ) \}}
\texttt{trunc} \ \texttt{fp_eval:n \{ round( \langle fpexpr_1 \rangle, \langle fpexpr_2 \rangle ) \}}
\texttt{floor} \ \texttt{fp_eval:n \{ round( \langle fpexpr_1 \rangle, \langle fpexpr_2 \rangle, \langle fpexpr_3 \rangle ) \}}

Only \texttt{round} accepts a third argument. Evaluates \( \langle \text{fpexpr}_1 \rangle = x \) and \( \langle \text{fpexpr}_2 \rangle = n \) and \( \langle \text{fpexpr}_3 \rangle = t \) then rounds \( x \) to \( n \) places. If \( n \) is an integer, this rounds \( x \) to a multiple of \( 10^{-n} \); if \( n = +\infty \), this always yields \( x \); if \( n = -\infty \), this yields one of \( \pm 0 \), \( \pm \infty \), or \text{NaN}; if \( n = \text{NaN} \), this yields \text{NaN}; if \( n \) is neither \( \pm \infty \) nor an integer, then an “invalid operation” exception is raised. When \( \langle \text{fpexpr}_3 \rangle \) is omitted, \( n = 0 \), i.e., \( \langle \text{fpexpr}_1 \rangle \) is rounded to an integer. The rounding direction depends on the function.

- \texttt{round} yields the multiple of \( 10^{-n} \) closest to \( x \), with ties (\( x \) half-way between two such multiples) rounded as follows. If \( t \) is \text{nan} (or not given) the even multiple is chosen (“ties to even”), if \( t = \pm 0 \) the multiple closest to \( 0 \) is chosen (“ties to zero”), if \( t \) is positive/negative the multiple closest to \( \infty/-\infty \) is chosen (“ties towards positive/negative infinity”).
- \texttt{floor} yields the largest multiple of \( 10^{-n} \) smaller or equal to \( x \) (“round towards negative infinity”);
- \texttt{ceil} yields the smallest multiple of \( 10^{-n} \) greater or equal to \( x \) (“round towards positive infinity”);
- \texttt{trunc} yields a multiple of \( 10^{-n} \) with the same sign as \( x \) and with the largest absolute value less than that of \( x \) (“round towards zero”).

“Overflow” occurs if \( x \) is finite and the result is infinite (this can only happen if \( \langle \text{fpexpr}_2 \rangle < -9984 \)). If any operand is a tuple, “invalid operation” occurs.

\texttt{sign} \ \texttt{fp_eval:n \{ sign( \langle fpexpr \rangle ) \}}

Evaluates the \( \langle \text{fpexpr} \rangle \) 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.
 Computes the sine, cosine, tangent, cotangent, cosecant, or secant of the \(\langle fpexpr \rangle\) given in radians. For arguments given in degrees, see \texttt{sind}, \texttt{cosd}, etc. Note that since \(\pi\) 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 \(\pm \infty\), 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.

 Computes the sine, cosine, tangent, cotangent, cosecant, or secant of the \(\langle fpexpr \rangle\) given in degrees. For arguments given in radians, see \texttt{sin}, \texttt{cos}, etc. Note that since \(\pi\) 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 \(\pm \infty\), 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.

 Computes the arcsine, arccosine, arccosecant, or arccscant of the \(\langle fpexpr \rangle\) and returns the result in radians, in the range \([-\pi/2, \pi/2]\) for \texttt{asin} and \texttt{acsc} and \([0, \pi]\) for \texttt{acos} and \texttt{asec}. For a result in degrees, use \texttt{asind}, \texttt{acosd}, etc. If the argument of \texttt{asin} or \texttt{acos} lies outside the range \([-1, 1]\), or the argument of \texttt{acsc} or \texttt{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.

 Computes the arcsine, arccosine, arccosecant, or arccscant of the \(\langle fpexpr \rangle\) and returns the result in degrees, in the range \([-90, 90]\) for \texttt{asin} and \texttt{acsc} and \([0, 180]\) for \texttt{acos} and \texttt{asec}. For a result in radians, use \texttt{asind}, \texttt{acosd}, etc. If the argument of \texttt{asin} or \texttt{acos} lies outside the range \([-1, 1]\), or the argument of \texttt{acsc} or \texttt{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.
.atan \mathit{fp}__eval: n \{ \mathit{atan}(\mathit{fpexpr}) \}\) 
.acot \mathit{fp}__eval: n \{ \mathit{acot}(\mathit{fpexpr}) \}\)
.acot \mathit{fp}__eval: n \{ \mathit{acot}(\mathit{fpexpr}_1,\mathit{fpexpr}_2) \}\)
.acot \mathit{fp}__eval: n \{ \mathit{acot}(\mathit{fpexpr}_1,\mathit{fpexpr}_2) \}\)

Those functions yield an angle in radians: \mathit{atand} and \mathit{acotd} are their analogs in degrees. The one-argument versions compute the arctangent or arccotangent of the \langle \mathit{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 \langle \mathit{fpexpr}_2,\mathit{fpexpr}_1 \rangle: this is the arctangent of \langle \mathit{fpexpr}_1/\mathit{fpexpr}_2 \rangle, possibly shifted by \pi depending on the signs of \langle \mathit{fpexpr}_1 \rangle and \langle \mathit{fpexpr}_2 \rangle. The two-argument arccotangent computes the angle in polar coordinates of the point \langle \mathit{fpexpr}_1,\mathit{fpexpr}_2 \rangle, equal to the arccotangent of \langle \mathit{fpexpr}_1/\mathit{fpexpr}_2 \rangle, possibly shifted by \pi. Both two-argument functions take values in the wider range \([-\pi,\pi]\). The ratio \langle \mathit{fpexpr}_1/\mathit{fpexpr}_2 \rangle need not be defined for the two-argument arctangent: when both expressions yield \pm 0, or when both yield \pm \infty, the resulting angle is one of \{\pm \pi/4,\pm 3\pi/4\} depending on signs. The “underflow” exception can occur. If any operand is a tuple, “invalid operation” occurs.

.atand \mathit{fp}__eval: n \{ \mathit{atand}(\mathit{fpexpr}) \}\)
.acotd \mathit{fp}__eval: n \{ \mathit{acotd}(\mathit{fpexpr}) \}\)
.acotd \mathit{fp}__eval: n \{ \mathit{acotd}(\mathit{fpexpr}_1,\mathit{fpexpr}_2) \}\)
.acotd \mathit{fp}__eval: n \{ \mathit{acotd}(\mathit{fpexpr}_1,\mathit{fpexpr}_2) \}\)

Those functions yield an angle in degrees: \mathit{atand} and \mathit{acotd} are their analogs in radians. The one-argument versions compute the arctangent or arccotangent of the \langle \mathit{fpexpr} \rangle: arctangent takes values in the range \([-90,90]\), and arccotangent in the range \([0,180]\). The two-argument arctangent computes the angle in polar coordinates of the point with Cartesian coordinates \langle \mathit{fpexpr}_2,\mathit{fpexpr}_1 \rangle: this is the arctangent of \langle \mathit{fpexpr}_1/\mathit{fpexpr}_2 \rangle, possibly shifted by 180 depending on the signs of \langle \mathit{fpexpr}_1 \rangle and \langle \mathit{fpexpr}_2 \rangle. The two-argument arccotangent computes the angle in polar coordinates of the point \langle \mathit{fpexpr}_1,\mathit{fpexpr}_2 \rangle, equal to the arccotangent of \langle \mathit{fpexpr}_1/\mathit{fpexpr}_2 \rangle, possibly shifted by 180. Both two-argument functions take values in the wider range \([-180,180]\). The ratio \langle \mathit{fpexpr}_1/\mathit{fpexpr}_2 \rangle need not be defined for the two-argument arctangent: when both expressions yield \pm 0, or when both yield \pm \infty, the resulting angle is one of \{\pm 45,\pm 135\} depending on signs. The “underflow” exception can occur. If any operand is a tuple, “invalid operation” occurs.

.sqrt \mathit{fp}__eval: n \{ \mathit{sqrt}(\mathit{fpexpr}) \}\)

Computes the square root of the \langle \mathit{fpexpr} \rangle. The “invalid operation” is raised when the \langle \mathit{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\{\text{rand()}\}

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 \texttt{\sys_rand_seed:} and set using \texttt{\sys_gset_rand_seed:n}.

\TeXhackers\note: This is based on pseudo-random numbers provided by the engine's primitive \texttt{\pdfuniformdeviate} in pdf\TeX, \texttt{\puniformdeviate} in \texttt{up\TeX} and \texttt{\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 \texttt{\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.

\fp_eval:n\{\text{randint(⟨fpexpr⟩)}\}
\fp_eval:n\{\text{randint(⟨fpexpr⟩₁, ⟨fpexpr⟩₂)}\}

Produces a pseudo-random integer between 1 and ⟨fpexpr⟩ or between ⟨fpexpr⟩₁ and ⟨fpexpr⟩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 \texttt{rand} for important comments on how these pseudo-random numbers are generated.

\texttt{inf}, \texttt{nan}

The special values $+\infty$, $-\infty$, and NaN are represented as \texttt{inf}, \texttt{-inf} and \texttt{nan} (see \texttt{\c_-inf_fp}, \texttt{\c_minus_inf_fp} and \texttt{\c_nan_fp}).

\texttt{pi}

The value of $\pi$ (see \texttt{\c_pi_fp}).

\texttt{deg}

The value of $1^\circ$ in radians (see \texttt{\c_one_degree_fp}).
Those units of measurement are equal to their values in pt, namely

- 1 in = 72.27 pt
- 1 pt = 1 pt
- 1 pc = 12 pt
- 1 cm = \( \frac{1}{25.4} \) in = 28.45275590551181 pt
- 1 mm = \( \frac{1}{25.4} \) in = 2.845275590551181 pt
- 1 dd = 0.376065 mm = 1.07000856496063 pt
- 1 cc = 12 dd = 12.84010277952756 pt
- 1 nd = 0.375 mm = 1.066978346456693 pt
- 1 nc = 12 nd = 12.80374015748031 pt
- 1 bp = \( \frac{1}{72} \) in = 1.00375 pt
- 1 sp = \( 2^{-16} \) pt = 1.52587890625 \times 10^{-5} pt.

The values of the (font-dependent) units \texttt{em} and \texttt{ex} are gathered from \TeX{} when the surrounding floating point expression is evaluated.

Other names for 1 and +0.

\begin{verbatim}
\fp_abs:n \{(floating point expression)\}
\end{verbatim}

Evaluates the \{(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 ±∞, NaN or a tuple, “invalid operation” occurs. Within floating point expressions, \texttt{abs()} can be used; it accepts ±∞ and NaN as arguments.

\begin{verbatim}
\fp_max:nn \{(fp expression 1) \} \{(fp expression 2)\}
\end{verbatim}

\begin{verbatim}
\fp_min:nn \{(fp expression 1) \} \{(fp expression 2)\}
\end{verbatim}

Evaluates the \{(floating point expressions)\} as described for \texttt{\fp_eval:n} and leaves the resulting larger (\texttt{max}) or smaller (\texttt{min}) 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{max()} and \texttt{min()} can be used.

## 28.10 Disclaimer and roadmap

The package may break down if the escape character is among 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.

- Function to count items in a tuple (and to determine if something is a tuple).
- Decide what exponent range to consider.

256
• Support signalling \texttt{nan}.
• Modulo and remainder, and rounding function \texttt{quantize} (and its friends analogous to \texttt{trunc}, \texttt{ceil}, \texttt{floor}).
• \texttt{\textbackslash fp\_format::nn \{(fpexpr)\} \{(format)\}}, but what should \texttt{(format)} be? More general pretty printing?
• 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}, \texttt{sinh}, \texttt{tanh}.
• Inverse hyperbolics.
• Base conversion, input such as \texttt{0xAB.CDEF}.
• Factorial (not with \texttt{!}), 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{!<>?}, \texttt{&\&}, \texttt{||}, and \texttt{!}.
• 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 $(-x) + x$ should depend on the rounding mode.
• \texttt{0e9999999999} gives a \texttt{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 28.9.1, write a grammar.
• It would be nice if the \verb|parse| auxiliaries for each operation were set up in the corresponding module, rather than centralizing in l3fp-parse.

• Some functions should get an \verb|_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 \verb|ternary| set of functions is ugly.

• There are many \verb|~| 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 \( c = \frac{2000}{\lfloor 200x \rfloor + 1} \in [10, 95] \) 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 (l3fp-basics).

• Understand and document \verb|\__fp_basics_pack_weird_low:NNNNw| and \verb|\__fp_basics_pack_weird_high:NNNNNNNNw| better. Move the other \verb|basics_pack| auxiliaries to 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 TeX fp packages, the international standards,\

• Also take into account the “inexact” exception?

• Support multi-character prefix operators (e.g., \verb|@/| or whatever)?
Chapter 29

The \texttt{l3fparray} package: fast global floating point arrays

29.1 \texttt{l3fparray} documentation

For applications requiring heavy use of floating points, this module provides arrays which can be accessed in constant time (contrast \texttt{l3seq}, where access time is linear). The interface is very close to that of \texttt{l3intarray}. The size of the array is fixed and must be given at point of initialisation.

\begin{verbatim}
\fparray_new:Nn \fparray new \{ \fparray \} \{ \langle \text{size} \rangle \}
\end{verbatim}

Evaluates the integer expression \langle \text{size} \rangle and allocates an \langle floating point array variable \rangle with that number of (zero) entries. The variable name should start with \texttt{\_g}_ because assignments are always global.

\begin{verbatim}
\fparray_count:N \fparray count \{ \fparray \}
\end{verbatim}

Expands to the number of entries in the \langle floating point array variable \rangle. This is performed in constant time.

\begin{verbatim}
\fparray_gset:Nnn \fparray get \{ \fparray \} \{ \langle \text{position} \rangle \} \{ \langle \text{value} \rangle \}
\end{verbatim}

Stores the result of evaluating the floating point expression \langle \text{value} \rangle into the \langle floating point array variable \rangle at the (integer expression) \langle \text{position} \rangle. If the \langle \text{position} \rangle is not between 1 and the \texttt{\fparray count}, an error occurs. Assignments are always global.

\begin{verbatim}
\fparray_gzero:N \fparray zero \{ \fparray \}
\end{verbatim}

Sets all entries of the \langle floating point array variable \rangle to +0. Assignments are always global.

\begin{verbatim}
\fparray_item:Nn \fparray item \{ \fparray \} \{ \langle \text{position} \rangle \}
\fparray_item_to_tl:Nn \fparray item to tl \{ \fparray \} \{ \langle \text{position} \rangle \}
\end{verbatim}

Applies \texttt{\fp use} or \texttt{\fp to tl} (respectively) to the floating point entry stored at the (integer expression) \langle \text{position} \rangle in the \langle floating point array variable \rangle. If the \langle \text{position} \rangle is not between 1 and the \texttt{\fparray count}, an error occurs.

259
Chapter 30

The \texttt{l3cctab} package

Category code tables

A category code table enables rapid switching of all category codes in one operation. For \LaTeX, this is possible over the entire Unicode range. For other engines, only the 8-bit range (0–255) is covered by such tables.

\section{Creating and initialising category code tables}

\begin{itemize}
\item \texttt{\cctab_new:N} \texttt{(category code table)}
\end{itemize}

\texttt{\cctab_new:N} creates a new \emph{category code table} variable or raises an error if the name is already taken. The declaration is global. The \emph{category code table} is initialised with the codes as used by \texttt{in\LaTeX}.

\begin{itemize}
\item \texttt{\cctab_const:Nn} \texttt{(category code table)} \texttt{\{category code set up\}}
\end{itemize}

\texttt{\cctab_const:Nn} creates a new \emph{category code table}, applies (in a group) the \emph{category code set up} on top of \texttt{in\LaTeX} settings, then saves them globally as a constant table. The \emph{category code set up} can include a call to \texttt{\cctab_select:N}.

\begin{itemize}
\item \texttt{\cctab_gset:Nn} \texttt{(category code table)} \texttt{\{category code set up\}}
\end{itemize}

\texttt{\cctab_gset:Nn} starting from the \texttt{in\LaTeX} category codes, applies (in a group) the \emph{category code set up}, then saves them globally in the \emph{category code table}. The \emph{category code set up} can include a call to \texttt{\cctab_select:N}.

\section{Using category code tables}

\begin{itemize}
\item \texttt{\cctab_begin:N} \texttt{(category code table)}
\end{itemize}

\texttt{\cctab_begin:N} switches locally the category codes in force to those stored in the \emph{category code table}. The prevailing codes before the function is called are added to a stack, for use with \texttt{\cctab_end:}. This function does not start a \TeX{} group.
\texttt{\cctab_end}:

Ends the scope of a \textit{category code table} started using \texttt{\cctab_begin:N}, returning the codes to those in force before the matching \texttt{\cctab_begin:N} was used. This must be used within the same \TeX{} group (and at the same \TeX{} group level) as the matching \texttt{\cctab_begin:N}.

\texttt{\cctab_select:N} \texttt{\cctab_select:c}

Selects the \textit{category code table} for the scope of the current group. This is in particular useful in the \textit{\setup} arguments of \texttt{\tl_set_rescan:Nnn}, \texttt{\tl_rescan:nn}, \texttt{\cctab_const:Nn}, and \texttt{\cctab_gset:Nn}.

\texttt{\cctab_item:Nn} \texttt{\cctab_item:cn} \texttt{\cctab_item:cn*}

New: 2021-05-10

\texttt{\cctab_if_exist_p:N} \texttt{\cctab_if_exist_p:c} \texttt{\cctab_if_exist:NTF} \texttt{\cctab_if_exist:cTF}

Tests whether the \textit{category code table} is currently defined. This does not check that the \textit{category code table} really is a category code table.

### 30.3 Category code table conditionals

### 30.4 Constant category code tables

\texttt{\c_code_cctab} \texttt{\c_document_cctab} \texttt{\c_initex_cctab} \texttt{\c_other_cctab} \texttt{\c_str_cctab}

Category code table for the expl3 code environment; this does \textit{not} include \texttt{ notch}, which is retained as an “other” character.

Category code table for a standard \LaTeX{} document, as set by the \LaTeX{} kernel. In particular, the upper-half of the 8-bit range will be set to “active” with \texttt{pdLaTeX} only. No \texttt{babel} shorthands will be activated.

Category code table as set up by ini\TeX{}.

Category code table where all characters have category code 12 (other).

Category code table where all characters have category code 12 (other) with the exception of spaces, which have category code 10 (space).
Part V
Text manipulation
Chapter 31

The \texttt{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.
Chapter 32

The \texttt{l3text} package: text processing

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 \langle text\rangle are normalized and become \{ and \}, respectively.

32.1 Expanding text

\texttt{text\_expand:n} \texttt{\{text\}}

Takes user input \langle text\rangle 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 \TeX{} protected are expanded exhaustively. Any commands listed in \l_text_expand_exclude_tl, \l_text_accenttl and \l_text_letterlike_tl are excluded from expansion.

\texttt{text\_declare\_expand\_equivalent:Nn} \texttt{\{replacement\}}

\texttt{text\_declare\_expand\_equivalent:cn}

Declares that the \langle replacement\rangle tokens should be used whenever the \langle cmd\rangle (a single token) is encountered. The \langle replacement\rangle tokens should be expandable.
32.2 Case changing

\text_lowercase:n \text_uppercase:n \text_titlecase:n \text_titlecase_first:n \text_lowercase:nn \text_uppercase:nn \text_titlecase:nn \text_titlecase_first:nn

Takes user input (text) first applies \text_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 (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 ij in Dutch which becomes IJ. The 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 \str_lowercase:n, \str_uppercase:n and \str_foldcase:n.

Case changing does not take place within math mode material so for example

\text_uppercase:n { Some-text-$y = mx + c$-with-{Braces} }

becomes

SOME TEXT $y = mx + c$ WITH {BRACES}

The arguments of commands listed in \_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 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 á can be case-changed using pdf\TeX{}. For \TeX{} 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 \textit{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 Eszett maps to a großes Eszett. Since there is a T1 slot for the großes Eszett 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 uppercasing; titlecasing leaves accents in place. (At present this is implemented only for Unicode engines.)

• 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 lowering of the relevant uppercase letters both when input as single Unicode codepoints and when using combining accents. The combining dot is removed when uppercasing 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 ij at the beginning of titlecased input produces IJ rather than IJ. The output retains two separate letters, thus this transformation is available using pdftex.

For titlecasing, note that there are two functions available. The function \textlowercase:n applies (broadly) uppercasing to the first letter of the input, then lowercasing to the remainder. In contrast, \texttitlecasefirst:n only carries out the uppercasing operation, and leaves the balance of the input unchanged. Determining whether non-letter characters at the start of text should switch from upper- to lowercasing is controllable. When \texttitlecasecheckletterbool is true, characters which are not letters (category code 11) are left unchanged and "skipped": the first letter is uppercased. (With 8-bit engines, this is extended to active characters which form part of a multi-byte letter codepoint.) When \texttitlecasecheckletterbool is false, the first character is uppercased, and the rest lowercased, irrespective of the nature of the character.

### 32.3 Removing formatting from text

\textpurify:n \{⟨text⟩\}

Takes user input ⟨text⟩ and expands as described for \textexpand:n, then removes all functions from the resulting text. Math mode material (as delimited by pairs given in \textmathdelims_tl or as the argument to commands listed in \textmatharg_tl) is left contained in a pair of $ delimiters. Non-expandable functions present in the ⟨text⟩ must either have a defined equivalent (see \textdeclarepurifyequivalent:Nn) or will be removed from the result. Implicit tokens are converted to their explicit equivalent.

\textdeclarepurifyequivalent:Nn \{⟨replacement⟩\}
\textdeclarepurifyequivalent:Nx

Declares that the ⟨replacement⟩ tokens should be used whenever the ⟨cmd⟩ (a single token) is encountered. The ⟨replacement⟩ tokens should be expandable.

### 32.4 Control variables

\textaccents_tl

Lists commands which represent accents, and which are left unchanged by expansion. (Defined only for the \LaTeX{} package.)
\l_text_letterlike_tl  Lists commands which represent letters; these are left unchanged by expansion. (Defined only for the \LaTeX{} 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_case_exclude_arg_tl  Lists commands which are excluded from case changing.

\l_text_expand_exclude_tl  Lists commands which are excluded from expansion.

\l_text_titlecase_check_letter_bool  Controls how the start of titlecasing is handled: when true, the first letter in text is considered. The standard setting is true.
Part VI
Typesetting
Chapter 33

The l3box package

Boxes

Box variables contain typeset material that can be inserted on the page or in other boxes. Their contents cannot be converted back to lists of tokens. There are three kinds of box operations: horizontal mode denoted with prefix \hbox_, vertical mode with prefix \vbox_, and the generic operations working in both modes with prefix \box_. For instance, a new box variable containing the words “Hello, world!” (in a horizontal box) can be obtained by the following code.

\box_new:N \l_hello_box
\hbox_set:Nn \l_hello_box { Hello, ~ world! }

The argument is typeset inside a TEX group so that any variables assigned during the construction of this box restores its value afterwards.

Box variables from l3box are compatible with those of L\LaTeX\2e and plain \TeX\ and can be used interchangeably. The l3box commands to construct boxes, such as \hbox:n or \hbox_set:Nn, are “color-safe”, meaning that

\hbox:n \{ \color_select:n \{ blue \} Hello, \} - world!

will result in “Hello,” taking the color blue, but “world!” remaining with the prevailing color outside the box.

33.1 Creating and initialising boxes

<table>
<thead>
<tr>
<th>\box_new:N</th>
<th>\box_new:N \langle box \rangle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creates a new \langle box \rangle or raises an error if the name is already taken. The declaration is global. The \langle box \rangle is initially void.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\box_clear:N</th>
<th>\box_clear:N \langle box \rangle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clears the content of the \langle box \rangle by setting the box equal to \c_empty_box.</td>
<td></td>
</tr>
</tbody>
</table>

269
Ensures that the ⟨box⟩ exists globally by applying \box_new:N if necessary, then applies \box_(g)clear:N to leave the ⟨box⟩ empty.

\box_set_eq:NN ⟨box1⟩ ⟨box2⟩
Sets the content of ⟨box1⟩ equal to that of ⟨box2⟩.

\box__if_exist_p:N * \box__if_exist_p:c * \box__if_exist:NTF * \box__if_exist:TTF *
Tests whether the ⟨box⟩ is currently defined. This does not check that the ⟨box⟩ really is a box.

33.2 Using boxes

\box_use:N ⟨box⟩
Inserts the current content of the ⟨box⟩ onto the current list for typesetting. An error is raised if the variable does not exist or if it is invalid.

\box_move_right:nn {dimexpr} {⟨box function⟩}
This function operates in vertical mode, and inserts the material specified by the ⟨box function⟩ such that its reference point is displaced horizontally by the given ⟨dimexpr⟩ from the reference point for typesetting, to the right or left as appropriate. The ⟨box function⟩ should be a box operation such as \box_use:N \langle box ⟩ or a “raw” box specification such as \vbox:n \{ xyz \}.

\box_move_up:nn {dimexpr} {⟨box function⟩}
This function operates in horizontal mode, and inserts the material specified by the ⟨box function⟩ such that its reference point is displaced vertically by the given ⟨dimexpr⟩ from the reference point for typesetting, up or down as appropriate. The ⟨box function⟩ should be a box operation such as \box_use:N \langle box ⟩ or a “raw” box specification such as \vbox:n \{ xyz \}.
33.3 Measuring and setting box dimensions

\box_dp:N  \box_dp:N \langle box \rangle
Calculates the depth (below the baseline) of the \langle box \rangle in a form suitable for use in a \langle dimension expression \rangle.

\textit{TEXhackers note:} This is the \TeX{} primitive \texttt{dp}.

\box_ht:N  \box_ht:N \langle box \rangle
Calculates the height (above the baseline) of the \langle box \rangle in a form suitable for use in a \langle dimension expression \rangle.

\textit{TEXhackers note:} This is the \TeX{} primitive \texttt{ht}.

\box_wd:N  \box_wd:N \langle box \rangle
Calculates the width of the \langle box \rangle in a form suitable for use in a \langle dimension expression \rangle.

\textit{TEXhackers note:} This is the \TeX{} primitive \texttt{wd}.

\box_ht_plus_dp:N  \box_ht_plus_dp:N \langle box \rangle
Calculates the total vertical size (height plus depth) of the \langle box \rangle in a form suitable for use in a \langle dimension expression \rangle.

\box_set_dp:Nn  \box_set_dp:cn \langle box \rangle \{(dimension expression)\}
Set the depth (below the baseline) of the \langle box \rangle to the value of the \{(dimension expression)\}.

\box_set_ht:Nn  \box_set_ht:cn \langle box \rangle \{(dimension expression)\}
Set the height (above the baseline) of the \langle box \rangle to the value of the \{(dimension expression)\}.

\box_set_wd:Nn  \box_set_wd:cn \langle box \rangle \{(dimension expression)\}
Set the width of the \langle box \rangle to the value of the \{(dimension expression)\}. 
33.4 Box conditionals

\box_if_empty_p:N \box_if_empty_p:c \box_if_empty:N \box_if_empty:cTF

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:N \box_if_horizontal:cTF

Tests if \texttt{<box>} is a horizontal box.

\box_if_vertical_p:N \box_if_vertical_p:c \box_if_vertical:N \box_if_vertical:cTF

Tests if \texttt{<box>} is a vertical box.

33.5 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 \texttt{(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.

33.6 Constant boxes

\c_empty_box

This is a permanently empty box, which is neither set as horizontal nor vertical.

\textbf{\TeXhackers note}: At the \TeX level this is a void box.

33.7 Scratch boxes

\l_tmpa_box \l_tmpb_box

Scratch boxes 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.

\g_tmpa_box \g_tmpb_box

Scratch boxes 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.
33.8 Viewing box contents

\box_show:N \box_show:c
Shows full details of the content of the \textit{box} in the terminal.

\box_show:Nnn \box_show:cn\n\n\textit{Updated: 2012-05-11}  
\eval{\textit{Updated: 2012-05-11}}
\n\eval{\textit{Updated: 2012-05-11}}
\eval{\textit{Updated: 2012-05-11}}
\eval{\textit{Updated: 2012-05-11}}

\box_log:N \box_log:c
Writes full details of the content of the \textit{box} to the log.

\box_log:Nnn \box_log:cn\n\n\textit{Updated: 2012-05-11}  
\eval{\textit{Updated: 2012-05-11}}
\n\eval{\textit{Updated: 2012-05-11}}
\eval{\textit{Updated: 2012-05-11}}
\eval{\textit{Updated: 2012-05-11}}
\eval{\textit{Updated: 2012-05-11}}

33.9 Boxes and color

All \LaTeX boxes are “color safe”: a color set inside the box stops applying after the end of the box has occurred.

33.10 Horizontal mode boxes

\hbox:n \hbox:cn\n\n\textit{Updated: 2017-04-05}  
\eval{\textit{Updated: 2017-04-05}}
\n\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}

\hbox_to_wd:nn \hbox_to_wd:cn\n\n\textit{Updated: 2017-04-05}  
\eval{\textit{Updated: 2017-04-05}}
\n\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}

\hbox_to_zero:n \hbox_to_zero:cn\n\n\textit{Updated: 2017-04-05}  
\eval{\textit{Updated: 2017-04-05}}
\n\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}

\hbox_set:Nn \hbox_set:cn \hbox_gset:Nn \hbox_gset:cn\n\n\textit{Updated: 2017-04-05}  
\eval{\textit{Updated: 2017-04-05}}
\n\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}
\eval{\textit{Updated: 2017-04-05}}

273
\hbox_set_to_wd:Nnn \hbox_set_to_wd:cn \hbox_gset_to_wd:Nnn \hbox_gset_to_wd:cn

Updated: 2017-04-05

\hbox_overlap_center:n \{\contents\}

Typesets the \{\contents\} into a horizontal box of zero width such that material protrudes equally to both sides of the insertion point.

\hbox_overlap_right:n \{\contents\}

Updated: 2017-04-05

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\}

Updated: 2017-04-05

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:

Updated: 2017-04-05

\hbox_set_to_wd:Nww \hbox_set_to_wd:cnw \hbox_gset_to_wd:Nww \hbox_gset_to_wd:cnw

Updated: 2017-06-08

\hbox_unpack:N \hbox_unpack:c

Unpacks the content of the horizontal \{box\}, retaining any stretching or shrinking applied when the \{box\} was set.

\TeX{} hackers note: This is the \TeX{} primitive \texttt{\unhcopy{}}.

33.11 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 \{(contents)\}
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 \{(contents)\}
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 \{(dimexpr)\} \{(contents)\}
Typesets the \langle contents \rangle into a vertical box of height \langle dimexpr \rangle and then includes this box in the current list for typesetting.

\vbox_to_zero:n \{(contents)\}
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 \{(contents)\}
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 \{(contents)\}
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 \{(dimexpr)\} \{(contents)\}
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 \langle box \rangle \langle contents \rangle \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

New: 2017-06-08

\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

33.12 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 } \hgroup_begin: \hbox_set:Nn \l_tmpa_box { B } \hgroup_begin: \box_use_drop:N \l_tmpa_box \hgroup_end: \hbox_show:N \l_tmpa_box \hgroup_end: \hbox_show:N \l_tmpa_box

the first use of \hbox_show:N will show an entirely cleared (void) box, and the second will show the letter 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 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.

\vbox_set_to_ht:Nnw \vbox_set_to_ht:NNn \vbox_set_to_ht:cnw \vbox_gset_to_ht:Nnw \vbox_gset_to_ht:cnw

Typesets the \textit{contents} to the height given by the \textit{dimexpr} and then stores the result inside the \textit{box}. In contrast to \vbox_set_to_ht:Nnn this function does not absorb the argument when finding the \textit{content}, and so can be used in circumstances where the \textit{content} 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)

Sets \textit{box1} to contain material to the height given by the \textit{dimexpr} by removing content from the top of \textit{box2} (which must be a vertical box).

\vbox_unpack:N \vbox_unpack:c

Unpacks the content of the vertical \textit{box}, retaining any stretching or shrinking applied when the \textit{box} was set.

\TeXhackers note: This is the \TeX primitive \texttt{unvcopy}.

276
\[\text{\box_use_drop:N}\]
\[\text{\box_use_drop:c}\]

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{\textit{T\TeX}hackers note:} This is the \texttt{\box} primitive.

\[\text{\box_set_eq_drop:NN}\]
\[\text{\box_gset_eq_drop:NN}\]

Sets the content of \langle box_1\rangle equal to that of \langle box_2\rangle, then drops \langle box_2\rangle.

\textbf{New: 2019-01-17}

\[\text{\hbox_unpack_drop:N}\]
\[\text{\vbox_unpack_drop:N}\]

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{\textit{T\TeX}hackers note:} This is the \texttt{T\TeX} primitive \texttt{\unhbox}.

\[\text{\box_gset_eq_drop:NN}\]

Sets the content of \langle box_1\rangle globally equal to that of \langle box_2\rangle, then drops \langle box_2\rangle.

\textbf{New: 2019-01-17}

33.13 Affine transformations

Affine transformations are changes which (informally) preserve straight lines. Simple translations are affine transformations, but are better handled in \texttt{T\TeX} 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 \langle x-size \rangle (horizontally) and \langle y-size \rangle (vertically); both of the sizes are dimension expressions. The \langle y-size \rangle 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 \{\langle x-size \rangle\} and \{\langle y-size \rangle\}, 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 \langle y-size \rangle results in the \langle box \rangle having a depth dependent on the height of the original and vice versa.

Resizes the \langle box \rangle to fit within the given \langle x-size \rangle (horizontally) and \langle y-size \rangle (vertically); both of the sizes are dimension expressions. The \langle y-size \rangle 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 \{\langle x-size \rangle\} and \{\langle y-size \rangle\}, 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 \langle y-size \rangle results in the \langle box \rangle having a depth dependent on the height of the original and vice versa.

Resizes the \langle box \rangle to \langle y-size \rangle (vertically), scaling the horizontal size by the same amount; \langle y-size \rangle is a dimension expression. The \langle y-size \rangle 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 \langle y-size \rangle 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 \langle y-size \rangle results in the \langle box \rangle having a depth dependent on the height of the original and vice versa.
Resizes the \{box\} to \{y-size\} (vertically), scaling the horizontal size by the same amount: \{y-size\} is a dimension expression. The \{y-size\} is the total vertical size (height plus depth). The updated \{box\} is an \texttt{hbox}, irrespective of the nature of the \{box\} before the resizing is applied. A negative \{y-size\} causes the material in the \{box\} to be reversed in direction, but the reference point of the \{box\} is unchanged. Thus a negative \{y-size\} results in the \{box\} having a depth dependent on the height of the original and \emph{vice versa}.

Resizes the \{box\} to \{x-size\} (horizontally), scaling the vertical size by the same amount: \{x-size\} is a dimension expression. The updated \{box\} is an \texttt{hbox}, irrespective of the nature of the \{box\} before the resizing is applied. A negative \{x-size\} causes the material in the \{box\} to be reversed in direction, but the reference point of the \{box\} is unchanged. Thus a negative \{x-size\} results in the \{box\} having a depth dependent on the height of the original and \emph{vice versa}.

Resizes the \{box\} to \{x-size\} (horizontally) and \{y-size\} (vertically): both of the sizes are dimension expressions. The \{y-size\} is the height only and does not include any depth. The updated \{box\} is an \texttt{hbox}, irrespective of the nature of the \{box\} before the resizing is applied. Negative sizes cause the material in the \{box\} to be reversed in direction, but the reference point of the \{box\} is unchanged. Thus a negative \{y-size\} results in the \{box\} having a depth dependent on the height of the original and \emph{vice versa}.

Resizes the \{box\} to \{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 \{box\} is an \texttt{hbox}, irrespective of the nature of the \{box\} before the resizing is applied. Negative sizes cause the material in the \{box\} to be reversed in direction, but the reference point of the \{box\} is unchanged. Thus a negative \{y-size\} results in the \{box\} having a depth dependent on the height of the original and \emph{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

Rotates the \textit{box} by \textit{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 \textit{box} is an \texttt{ hbox}, irrespective of the nature of the \textit{box} before the rotation is applied.

Scales the \textit{box} by factors \textit{x-scale} and \textit{y-scale} in the horizontal and vertical directions, respectively (both scales are integer expressions). The updated \textit{box} is an \texttt{ hbox}, irrespective of the nature of the \textit{box} before the scaling is applied. Negative scalings cause the material in the \textit{box} to be reversed in direction, but the reference point of the \textit{box} is unchanged. Thus a negative \textit{y-scale} results in the \textit{box} having a depth dependent on the height of the original and \textit{vice versa}.

### 33.14 Primitive box conditionals

\if_hbox:N \if_hbox:N  \texttt{if}
\begin{verbatim}
\if_hbox:N (box) \{true code\}
  \else:
    \{false code\}
\fi:
\end{verbatim}
Tests is \textit{box} is a horizontal box.

\textbf{\TeXhackers note:} This is the \TeX\ primitive \texttt{ifhbox}.

\if_vbox:N \if_vbox:N \texttt{if}
\begin{verbatim}
\if_vbox:N (box) \{true code\}
  \else:
    \{false code\}
\fi:
\end{verbatim}
Tests is \textit{box} is a vertical box.

\textbf{\TeXhackers note:} This is the \TeX\ primitive \texttt{ifvbox}.

\if_box_empty:N \if_box_empty:N \texttt{if}
\begin{verbatim}
\if_box_empty:N (box) \{true code\}
  \else:
    \{false code\}
\fi:
\end{verbatim}
Tests is \textit{box} is an empty (void) box.

\textbf{\TeXhackers note:} This is the \TeX\ primitive \texttt{ifvoid}.

280
Chapter 34

The \texttt{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 \texttt{xcoffins} module (in the \texttt{l3experimental} bundle).

34.1 Creating and initialising coffins

\begin{verbatim}
\coffin_new:N \coffin_new:c
\coffin_clear:N \coffin_clear:c
\coffin_gclear:N \coffin_gclear:c
\coffin_set_eq:NN \coffin_set_eq:NN \coffin_gset_eq:NN \coffin_gset_eq:NN
\coffin_if_exist_p:N \coffin_if_exist_p:c \coffin_if_exist:NTF \coffin_if_exist:CTF
\end{verbatim}

\texttt{\coffin_new:N} \texttt{\coffin_new:c}

Creates a new \texttt{\coffin} or raises an error if the name is already taken. The declaration is global. The \texttt{\coffin} is initially empty.

\texttt{\coffin_clear:N} \texttt{\coffin_clear:c}

Clears the content of the \texttt{\coffin}.

\texttt{\coffin_set_eq:NN} \texttt{\coffin_set_eq:NN} \texttt{\coffin_gset_eq:NN} \texttt{\coffin_gset_eq:NN}

Sets both the content and poles of \texttt{\coffin_1} equal to those of \texttt{\coffin_2}.

\texttt{\coffin_if_exist_p:N} \texttt{\coffin_if_exist_p:c} \texttt{\coffin_if_exist:NTF} \texttt{\coffin_if_exist:CTF}

Tests whether the \texttt{\coffin} is currently defined.
34.2 Setting coffin content and poles

\hcoffin_set:Nn \hcoffin_set:cn \hcoffin_gset:Nn \hcoffin_gset:cn

\hcoffin_set:Nw \hcoffin_set:cw \hcoffin_gset:Nw \hcoffin_gset:cw

\vcoffin_set:Nnn \vcoffin_set:cn \vcoffin_gset:Nnn \vcoffin_gset:cn

\coffin_set_horizontal_pole:Nnn \coffin_set_horizontal_pole:cn \coffin_gset_horizontal_pole:Nnn \coffin_gset_horizontal_pole:cn

Sets the \textit{pole} to run horizontally through the \textit{coffin}. The \textit{pole} is placed at the \textit{offset} from the bottom edge of the bounding box of the \textit{coffin}. The \textit{offset} should be given as a dimension expression.
Sets the \(\textit{pole}\) to run vertically through the \(\textit{coffin}\). The \(\textit{pole}\) is placed at the \(\textit{offset}\) from the left-hand edge of the bounding box of the \(\textit{coffin}\). The \(\textit{offset}\) should be given as a dimension expression.

### 34.3 Coffin affine transformations

Resized the \(\textit{coffin}\) to \(\textit{width}\) and \(\textit{total-height}\), both of which should be given as dimension expressions.

Rotates the \(\textit{coffin}\) by the given \(\textit{angle}\) (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.

Scales the \(\textit{coffin}\) by a factors \(\textit{x-scale}\) and \(\textit{y-scale}\) in the horizontal and vertical directions, respectively. The two scale factors should be given as real numbers.

### 34.4 Joining and using coffins

This function attaches \(\textit{coffin}_2\) to \(\textit{coffin}_1\) such that the bounding box of \(\textit{coffin}_1\) is not altered, i.e. \(\textit{coffin}_2\) can protrude outside of the bounding box of the coffin. The alignment is carried out by first calculating \(\textit{handle}_1\), the point of intersection of \(\textit{coffin}_1\)-pole\_1 and \(\textit{coffin}_1\)-pole\_2, and \(\textit{handle}_2\), the point of intersection of \(\textit{coffin}_2\)-pole\_1 and \(\textit{coffin}_2\)-pole\_2. \(\textit{coffin}_2\) is then attached to \(\textit{coffin}_1\) such that the relationship between \(\textit{handle}_1\) and \(\textit{handle}_2\) is described by the \(\textit{x-offset}\) and \(\textit{y-offset}\). The two offsets should be given as dimension expressions.
This function joins \texttt{⟨coffin1⟩} to \texttt{⟨coffin2⟩} 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 \texttt{⟨handle⟩}, the point of intersection of \texttt{⟨pole1⟩} and \texttt{⟨pole2⟩}. The coffin is then typeset in horizontal mode such that the relationship between the current reference point in the document and the \texttt{⟨handle⟩} is described by the \texttt{⟨x-offset⟩} and \texttt{⟨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.

34.5 Measuring coffins

\texttt{\coffin_dp:N \coffin_dp:c}

Calculates the depth (below the baseline) of the \texttt{⟨coffin⟩} in a form suitable for use in a \texttt{⟨dimension expression⟩}.

\texttt{\coffin_ht:N \coffin_ht:c}

Calculates the height (above the baseline) of the \texttt{⟨coffin⟩} in a form suitable for use in a \texttt{⟨dimension expression⟩}.

\texttt{\coffin_wd:N \coffin_wd:c}

Calculates the width of the \texttt{⟨coffin⟩} in a form suitable for use in a \texttt{⟨dimension expression⟩}.

34.6 Coffin diagnostics

\texttt{\coffin_display_handles:Nn \coffin_display_handles:cn}

This function first calculates the intersections between all of the \texttt{⟨poles⟩} of the \texttt{⟨coffin⟩} to give a set of \texttt{⟨handles⟩}. It then prints the \texttt{⟨coffin⟩} at the current location in the source, with the position of the \texttt{⟨handles⟩} marked on the coffin. The \texttt{⟨handles⟩} are labelled as part of this process: the locations of the \texttt{⟨handles⟩} and the labels are both printed in the \texttt{⟨color⟩} specified.
This function first calculates the ⟨handle⟩ for the ⟨coffin⟩ as defined by the intersection of ⟨pole1⟩ and ⟨pole2⟩. It then marks the position of the ⟨handle⟩ on the ⟨coffin⟩. The ⟨handle⟩ are labelled as part of this process: the location of the ⟨handle⟩ and the label are both printed in the ⟨color⟩ specified.

This function shows the structural information about the ⟨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 latter, rather than the absolute values, which determines the direction of the pole.

This function writes the structural information about the ⟨coffin⟩ in the log file. See also \texttt{\coffin
show_structure:N} which displays the result in the terminal.

Shows full details of poles and contents of the ⟨coffin⟩ in the terminal or log file. See \texttt{\coffin
show_structure:N} and \texttt{\box
show:N} to show separately the pole structure and the contents.

Shows poles and contents of the ⟨coffin⟩ in the terminal or log file, showing the first \(\langle\text{intexpr}\rangle\) items in the coffin, and descending into \(\langle\text{intexpr}\rangle\) group levels. See \texttt{\coffin
_show_structure:N} and \texttt{\box
show:N} to show separately the pole structure and the contents.

34.7 Constants and variables

A permanently empty 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.
\g_tmpa_coffin \g_tmpb_coffin

New: 2019-01-24

Scratch coffins 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.
Chapter 35

The l3color package
Color support

35.1 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.

\color_group_begin:\color_group_end:
New: 2011-09-03

Creates a color group: one used to “trap” color settings. This grouping is built in to for example \hbox_set:Nn.

\color_ensure_current:
New: 2011-09-03

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 \color_group_begin: ... \color_group_end: group.

35.2 Color models

A color model is a way to represent sets of colors. Different models are particularly suitable for different output methods, e.g. screen or print. Parameter-based models can describe a very large number of unique colors, and have a varying number of axes which define a color space. In contrast, various proprietary models are available which define spot colors (more formally separations).

Core models are used to pass color information to output; these are “native” to l3color. Core models use real numbers in the range [0,1] to represent values. The core models supported here are

- gray Grayscale color, with a single axis running from 0 (fully black) to 1 (fully white)
- rgb Red-green-blue color, with three axes, one for each of the components
• **cmyk** Cyan-magenta-yellow-black color, with four axes, one for each of the components.

There are also interface models: these are convenient for users but have to be manipulated before storing/passing to the backend. Interface models are primarily integer-based: see below for more detail. The supported interface models are

• **Gray** Grayscale color, with a single axis running from 0 (fully black) to 15 (fully white)

• **hsb** Hue-saturation-brightness color, with three axes, all real values in the range $[0, 1]$ for hue saturation and brightness

• **Hsb** Hue-saturation-brightness color, with three axes, integer in the range $[0, 360]$ for hue, real values in the range $[0, 1]$ for saturation and brightness

• **HSB** Hue-saturation-brightness color, with three axes, integers in the range $[0, 240]$ for hue, saturation and brightness

• **HTML** HTML format representation of RGB color given as a single six-digit hexadecimal number

• **RGB** Red-green-blue color, with three axes, one for each of the components, values as integers from 0 to 255

• **wave** Light wavelength, a real number in the range 380 to 780 (nanometres)

All interface models are internally stored as **rgb**.

To allow parsing of data from **xcolor**, any leading model up the first : will be discarded; the approach of selecting an internal form for data is not used in **l3color**.

Additional models may be created to allow mixing of separation colors with each other or with those from other models. See Section 35.9 for more detail of color support for additional models.

When color is selected by model, the (*values*) given are specified as a comma-separated list. The length of the list will therefore be determined by the detail of the model involved.

Color models (and interconversion) are complex, and more details are given in the manual to the **l3tex2e xcolor** package and in the *PostScript Language Reference Manual*, published by Addison–Wesley.

### 35.3 Color expressions

In addition to allowing specification of color by model and values, **l3color** also supports color expressions. These are created by combining one or more color names, with the amount of each specified as a percentage. The latter is given between ! symbols in the expression. Thus for example

```
red!50!green
```

is a mixture of 50% red and 50% green. A trailing percentage is interpreted as implicitly followed by **white**, and so

```
red!25
```
specifies 25\% red mixed with 75\% white.

Where the models for the mixed colors are different, the model of the first color is used. Thus

\texttt{red!50!cyan}

will result in a color specification using the \texttt{rgb} model, made up of 50\% red and 50\% of cyan \textit{expressed in rgb}. This may be important as color model interconversion is not exact.

The one exception to the above is where the first model in an expression is \texttt{gray}. In this case, the order of mixing is “swapped” internally, so that for example

\texttt{black!50!red}

has the same result as

\texttt{red!50!black}

(the predefined colors \texttt{black} and \texttt{white} use the \texttt{gray} model).

Where more than two colors are mixed in an expression, evaluation takes place in a stepwise fashion. Thus in

\texttt{cyan!50!magenta!10!yellow}

the sub-expression

\texttt{cyan!50!magenta}

is first evaluated to give an intermediate color specification, before the second step

\texttt{<intermediate>!10!yellow}

where \texttt{<intermediate>} represents this transitory calculated value.

Within a color expression, . may be used to represent the color active for typesetting (the current color). This allows for example

\texttt{.!50}

to mean a mixture of 50\% of current color with white.

(Color expressions supported here are a subset of those provided by the \LaTeX \texttt{2e xcolor} package. At present, only such features as are clearly useful have been added here.)

### 35.4 Named colors

Color names are stored in a single namespace, which makes them accessible as part of color expressions. Whilst they are not reserved in a technical sense, the names \texttt{black}, \texttt{white}, \texttt{red}, \texttt{green}, \texttt{blue}, \texttt{cyan}, \texttt{magenta} and \texttt{yellow} have special meaning and should not be redefined. Color names should be made up of letters, numbers and spaces only: other characters are reserved for use in color expressions. In particular, . represents the current color at the start of a color expression.

\begin{verbatim}
\color_set:nn \{name\} \{color expression\}
\end{verbatim}

Evaluates the \texttt{(color expression)} and stores the resulting color specification as the \texttt{(name)}.

289
\color_set:nn \{\text{name}\} \{\text{model(s)}\} \{\text{value(s)}\}
Stores the color specification equivalent to the \text{\textlangle model\textrangle} and \text{\textlangle value\textrangle} as the \text{\textlangle name\textrangle}.

\color_set_eq:nn \{\text{name1}\} \{\text{name2}\}
Copies the color specification in \text{\textlangle name2\textrangle} to \text{\textlangle name1\textrangle}. The special name . may be used to represent the current color, allowing it to be saved to a name.

\color_show:n \color_log:n
\color_log:n \{\text{name}\}
Displays the color specification stored in the \text{\textlangle name\textrangle} on the terminal or log file.

35.5 Selecting colors

General selection of color is safe when split across pages: a stack is used to ensure that the correct color is re-selected on the new page.

\color_select:n \color_select:nn \{\text{color expression}\}
Parses the \text{\textlangle color expression\textrangle} and then activates the resulting color specification for type-set material.

\color_fill:n \color_fill:nn \{\text{model(s)}\} \{\text{value(s)}\}
Activates the color specification equivalent to the \text{\textlangle model\textrangle} and \text{\textlangle value\textrangle} for type-set material.

\l_color_fixed_model_tl
When this is set to a non-empty value, colors will be converted to the specified model when they are selected. Note that included images and similar are not influenced by this setting.

35.6 Colors for fills and strokes

Colors for drawing operations and so forth are split into strokes and fills (the latter may also be referred to as non-stroke color). The fill color is used for text under normal circumstances. Depending on the backend, stroke color may use a stack, in which case it exhibits the same page breaking behavior as general color. However, \texttt{dvips/dvisvgm} do not support this, and so color will need to be contained within a scope, such as \texttt{\draw_begin::\draw_end::}. Note that the current color is the fill color, as this is used for running text.

\color_fill:n \color_stroke:n \color_fill:nn \color_stroke:nn
Parses the \text{\textlangle color expression\textrangle} and then activates the resulting color specification for filling or stroking.

Activates the color specification equivalent to the \text{\textlangle model\textrangle} and \text{\textlangle value\textrangle} for filling or stroking.
When using `dvips`, this PostScript variables hold the stroke color.

### 35.7 Multiple color models

When selecting or setting a color with an explicit model, it is possible to give values for more than one model at one time. This is particularly useful where automated conversion between models does not give the desired outcome. To do this, the list of models and list of values are both subdivided using `/` characters (as for the similar function in `xcolor`). For example, to save a color with explicit `cmyk` and `rgb` values, one could use

```
\color_set:nnn { foo } { cmyk / rgb }
{ 0.1 , 0.2 , 0.3 , 0.4 / 0.1, 0.2 , 0.3 }
```

The manually-specified conversion will be used in preference to automated calculation whenever the model(s) listed are used: both in expressions and when a fixed model is active.

Similarly, the same syntax can be applied to directly selecting a color.

```
\color_select:nn { cmyk / rgb }
{ 0.1 , 0.2 , 0.3 , 0.4 / 0.1, 0.2 , 0.3 }
```

Again, this list is used when a fixed model is active: the first entry is used unless there is a fixed model matching one of the other entries.

### 35.8 Exporting color specifications

The major use of color expressions is in setting typesetting output, but there are other places in which some form of color information is required. These may need data in a different format or using a different model to the internal representation. Thus a set of functions are available to export colors in different formats.

Valid export targets are

- **backend** Two brace groups: the first containing the model, the second containing space-separated values appropriate for the model; this is the format required by backend functions of `expl3`
- **comma-sep-cmyk** Comma-separated cyan-magenta-yellow-black values
- **comma-sep-rgb** Comma-separated red-green-blue values suitable for use as a PDF annotation color
- **HTML** Uppercase two-digit hexadecimal values, expressing a red-green-blue color; the digits are *not* separated
- **space-sep-cmyk** Space-separated cyan-magenta-yellow-black values
- **space-sep-rgb** Space-separated red-green-blue values suitable for use as a PDF annotation color
Parses the \( \langle \text{color expression} \rangle \) as described earlier, then converts to the \( \langle \text{format} \rangle \) specified and assigns the data to the \( \langle \text{tl} \rangle \).

Expresses the combination of \( \langle \text{model} \rangle \) and \( \langle \text{value(s)} \rangle \) in an internal representation, then converts to the \( \langle \text{format} \rangle \) specified and assigns the data to the \( \langle \text{tl} \rangle \).

### 35.9 Creating new color models

Additional color models are required to support specialist workflows, for example those involving separations (see [https://helpx.adobe.com/indesign/using/spot-process-colors.html](https://helpx.adobe.com/indesign/using/spot-process-colors.html) for details of the use of separations in print). Color models may be split into families; for the standard device-based color models (DeviceCMYK, DeviceRGB, DeviceGray), these are synonymous. This is not generally the case: see the PDF reference for more details. (Note that l3color uses the shorter names cmyk, etc.)

For a Separation space, there are three compulsory keys.

- **name** The name of the Separation, for example the formal name of a spot color ink. Such a \( \langle \text{name} \rangle \) may contain spaces, etc., which are not permitted in the \( \langle \text{model} \rangle \).
- **alternative-model** An alternative device colorspace, one of cmyk, rgb, gray or CIELAB. The three parameter-based models work as described above; see below for details of CIELAB colors.
- **alternative-values** A comma-separated list of values appropriate to the \( \langle \text{alternative-model} \rangle \). This information is used by the PDF application if the Separation is not available.

CIELAB color separations are created using the \( \langle \text{alternative-model} = \text{CIELAB} \rangle \) setting. These colors must also have an \( \langle \text{illuminant} \rangle \) key, one of a, c, e, d50, d55, d65 or d75. The \( \langle \text{alternative-values} \rangle \) in this case are the three parameters \( L^* \), \( a^* \) and \( b^* \) of the CIELAB model. Full details of this device-independent color approach are given in the documentation to the colorspace package.

CIELAB colors cannot be converted into other device-dependent color spaces, and as such, mixing can only occur if colors set up using the CIELAB model are also given with an alternative parameter-based model. If that is not the case, l3color will fallback to using black as the colorant in any mixing.

For a DeviceN space, there is one compulsory key.
• **names** The names of the components of the DeviceN space. Each should be either the (name) of a Separation model, a process color name (cyan, etc.) or the special name none.
Chapter 36

The l3pdf package
Core PDF support

36.1 Objects

\pdf_object_new:nn \pdf_object_new:nn \{object\} \{type\}
\pdf_object_new:nn \{object\} \{type\}
\pdf_object_new:nn \{object\} \{type\}
Declares \{object\} as a PDF object of \{type\}, which should be one of
- array
- dict
- fstream
- stream
The object may be referenced from this point on, and written later using \pdf_object_write:nn.

\pdf_object_if_exist_p:n \pdf_object_if_exist_p:n \{object\}
\pdf_object_if_exist_p:n \{object\}
\pdf_object_if_exist_p:n \{object\}
Tests whether an object with name \{object\} has been defined.

\pdf_object_write:nn \pdf_object_write:nn \{object\} \{content\}
\pdf_object_write:nn \pdf_object_write:nn \{object\} \{content\}
\pdf_object_write:nn \pdf_object_write:nn \{object\} \{content\}
\pdf_object_write:nn \pdf_object_write:nx
\pdf_object_write:nx \{object\} \{content\}
\pdf_object_write:nx \{object\} \{content\}
\pdf_object_write:nx \{object\} \{content\}
Writes the \{content\} as content of the \{object\}. Depending on the \{type\} declared for the object, the format required for the \{data\} will vary
array A space-separated list of values
dict Key–value pairs in the form /\{key\} \{value\}
fstream Two brace groups: \{file name\} and \{file content\}
stream Two brace groups: \{attributes (dictionary)\} and \{stream contents\}
\pdf_object_ref:n \pdf_object_ref:n \{object\}

Inserts the appropriate information to reference the \{object\} in, for example, page resource allocation.

\pdf_object_unnamed_write:nn \pdf_object_unnamed_write:nn \{type\} \{content\}
\pdf_object_unnamed_write:nx

Writes \{content\} as content of an anonymous object. Depending on the \{type\}, the format required for the \{data\} will vary:

array A space-separated list of values

dict Key–value pairs in the form /\{key\} \{value\}

fstream Two brace groups: \{attributes (dictionary)\} and \{file name\}

stream Two brace groups: \{attributes (dictionary)\} and \{stream contents\}

\pdf_object_ref_last: \pdf_object_ref_last:

Inserts the appropriate information to reference the last \{object\} created. This is particularly useful for anonymous objects.

\pdf_pageobject_ref:n \pdf_pageobject_ref:n \{pageobject\}

Inserts the appropriate information to reference the \{pageobject\}.

### 36.2 Version

\pdf_version_compare_p:Nn \pdf_version_compare_p:NnTF \{comparator\} \{\{version\}\} \{\{true code\}\} \{\{false code\}\}
\pdf_version_compare_p:NnTF

Compares the version of the PDF being created with the \{version\} string specified, using the \{comparator\}. Either the \{true code\} or \{false code\} will be left in the output stream.

\pdf_version_gset:n \pdf_version_gset:n \{version\}
\pdf_version_min_gset:n

Sets the \{version\} of the PDF being created. The min version will not alter the output version unless it is currently lower than the \{version\} requested. This function may only be used up to the point where the PDF file is initialised. With dvips it sets \pdf_version_major: and \pdf_version_minor: and allows to compare the values with \pdf_version_compare:Nn, but the PDF version itself still has to be set with the command line option -dCompatibilityLevel of ps2pdf.

\pdf_version: \pdf_version: \pdf_version_major: \pdf_version_minor:

Expands to the currently-active PDF version.

295
36.3 Compression

\texttt{\textbackslash pdf\_uncompress:}

Disables any compression of the PDF, where possible.

This function may only be used up to the point where the PDF file is initialised.

36.4 Destinations

Destinations are the places a link jumped too. Unlike the name may suggest they don’t described an exact location in the PDF. Instead a destination contains a reference to a page along with an instruction how to display this page. The normally used “XYZ top left zoom” for example instructs the viewer to show the page with the given zoom and the top left corner at the top left coordinates—which then gives the impression that there is an anchor at this position.

If an instruction takes a coordinate, it is calculated by the following commands relative to the location the command is issued. So to get a specific coordinate one has to move the command to the right place.

\texttt{\textbackslash pdf\_destination:nn} \{\langle name\rangle\} \{(type or integer)\}

This creates a destination. \{(type or integer)\} can be one of \texttt{fit}, \texttt{fith}, \texttt{fitv}, \texttt{fitb}, \texttt{fitbh}, \texttt{fitbv}, \texttt{fitr}, \texttt{xyz} or an integer representing a scale factor in percent. \texttt{fitr} here gives only a lightweight version of /FitR: The backend code defines \texttt{fitr} so that it will with \texttt{pdflatex} and \texttt{ LuaLaTeX} use the coordinates of the surrounding box, with \texttt{dvips} and \texttt{dvipdfmx} it falls back to \texttt{fit}. For full control use \texttt{\pdf_destination:nnnn}.

The keywords match to the PDF names as described in the following tabular.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>PDF</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>fit</td>
<td>/Fit</td>
<td>Fits the page to the window</td>
</tr>
<tr>
<td>fith</td>
<td>/FitH top</td>
<td>Fits the width of the page to the window</td>
</tr>
<tr>
<td>fitv</td>
<td>/FitV left</td>
<td>Fits the height of the page to the window</td>
</tr>
<tr>
<td>fitb</td>
<td>/FitB</td>
<td>Fits the page bounding box to the window</td>
</tr>
<tr>
<td>fitbh</td>
<td>/FitBH top</td>
<td>Fits the width of the page bounding box to the window.</td>
</tr>
<tr>
<td>fitbv</td>
<td>/FitBV left</td>
<td>Fits the height of the page bounding box to the window.</td>
</tr>
<tr>
<td>fitr</td>
<td>/FitR left bottom right top</td>
<td>Fits the rectangle specified by the four coordinates to the window (see above for the restrictions)</td>
</tr>
<tr>
<td>xyz</td>
<td>/XYZ left top null</td>
<td>Sets a coordinate but doesn’t change the zoom</td>
</tr>
<tr>
<td>{(integer)}</td>
<td>/XYZ left top zoom</td>
<td>Sets a coordinate and a zoom meaning {(integer)}%.</td>
</tr>
</tbody>
</table>
\pdf_destination:nnnn \pdf_destination:nnnn \{name\} \{width\} \{height\} \{depth\}

This creates a destination with /FitR type with the given dimensions relative to the current location. The destination is in a box of size zero, but it doesn’t switch to horizontal mode.
Part VII
Additions and removals
Chapter 37

The l3candidates package
Experimental additions to l3kernel

37.1 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!
37.2 Additions to l3box

\box_clip:N \{box\}

Clips the \{box\} in the output so that only material inside the bounding box is displayed in the output. The updated \{box\} is an hbox, irrespective of the nature of the \{box\} before the clipping is applied.

These functions require the \LaTeXX native drivers: they do not work with the \LaTeXX graphics drivers!

TeXhackers 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\} \{\left\} \{\bottom\} \{\right\} \{\top\}

Adjusts the bounding box of the \{box\} \{\left\} is removed from the left-hand edge of the bounding box, \{\right\} from the right-hand edge and so forth. All adjustments are \{dimension expressions\}. Material outside of the bounding box is still displayed in the output unless \box_clip:N is subsequently applied. The updated \{box\} is an hbox, irrespective of the nature of the \{box\} 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\} \{\llx\} \{\lly\} \{\urx\} \{\ury\}

Adjusts the bounding box of the \{box\} such that it has lower-left co-ordinates \{\llx, \lly\} and upper-right co-ordinates \{\urx, \ury\}. All four co-ordinate positions are \{dimension expressions\}. Material outside of the bounding box is still displayed in the output unless \box_clip:N is subsequently applied. The updated \{box\} is an hbox, irrespective of the nature of the \{box\} before the viewport operation is applied.

37.3 Additions to l3expan

\exp_args_generate:n \{variant argument specifiers\}

Defines \exp_args:N variant functions for each \{variant\} given in the comma list \{\{variant argument specifiers\}\}. Each \{variant\} 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 \{variant\}. This is only useful for cases where \cs_generate_variant:Nn is not applicable.

37.4 Additions to l3fp

\fp_if_nan:n \{fpexpr\}

Evaluates the \{fpexpr\} and tests whether the result is exactly \texttt{NaN}. The test returns \texttt{false} for any other result, even a tuple containing \texttt{NaN}. 
37.5 Additions to l3file

\iow_allow_break:
\iow_allow_break:
New: 2018-12-29
In the first argument of \iow_wrap:nnN (for instance in messages), inserts a break-point
that allows a line break. In other words this is a zero-width breaking space.

\ior_get_term:nN \ior_get_term:nN \langle prompt \rangle \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. \iow_term:n.
Where the \langle prompt \rangle is given, it will appear in the terminal followed by an \texttt{=}, e.g.
prompt=

\ior_shell_open:Nn \ior_shell_open:Nn \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:nnNTF.

37.6 Additions to l3flag

\flag_raise_if_clear:n \flag_raise_if_clear:n \langle flag name \rangle
Ensures the \langle flag \rangle is raised by making its height at least 1, locally.

37.7 Additions to l3intarray

\intarray_gset_rand:Nnn \intarray_gset_rand:Nnn \langle intarray var \rangle \{ \langle minimum \rangle \} \{ \langle maximum \rangle \}
\intarray_gset_rand:Nn \intarray_gset_rand:Nn \langle intarray var \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.

\intarray_to_clist:N \intarray_to_clist:N \langle intarray var \rangle
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.

301
37.8 Additions to \texttt{l3msg}

\begin{verbatim}
\msg_show_eval:Nn \msg_log_eval:Nn
Rem: 2017-12-04
\end{verbatim}

Shows or logs the \texttt{(expression)} (turned into a string), an equal sign, and the result of applying the \texttt{(function)} to the \texttt{(expression)} (with \texttt{f}-expansion). For instance, if the \texttt{(function)} is \texttt{\int_eval:n} and the \texttt{(expression)} is \texttt{1+2} then this logs \texttt{> 1+2=3}.

\begin{verbatim}
\msg_show_item:n * \seq_map_function:NN \msg_show_item:n
\msg_show_item_unbraced:n * \prop_map_function:NN \msg_show_item:nn
\msg_show_item:nn * \msg_show_item_unbraced:nn *
Rem: 2017-12-04
\end{verbatim}

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.

37.9 Additions to \texttt{l3prg}

\begin{verbatim}
\bool_set_inverse:N \bool_set_inverse:c
\bool_gset_inverse:N \bool_gset_inverse:c
Rem: 2018-05-10
\end{verbatim}

\texttt{\bool_set_inverse:N} (\texttt{boolean})

Toggles the \texttt{(boolean)} from \texttt{true} to \texttt{false} and conversely: sets it to the inverse of its current value.
\bool_case_true:n  \bool_case_true:nTF
\bool_case_true:nTF  \bool_case_true:nTF
\bool_case_false:n \bool_case_false:nTF
\bool_case_false:nTF

Evaluates in turn each of the \emph{boolean expression cases} 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 \}
\{ No idea! \}
\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 ... elseif ... elseif ... else ...”.

37.10 Additions to \texttt{l3prop}

\prop_rand_key_value:N  \prop_rand_key_value:N \prop_rand_key_value:N \prop_rand_key_value:N \prop_rand_key_value:N \prop_rand_key_value:N \prop_rand_key_value:N
\prop_rand_key_value:c  \prop_rand_key_value:c

Selects a pseudo-random key–value pair from the \texttt{\langle property list \rangle} and returns \texttt{\langle key \rangle} and \texttt{\langle value \rangle}. If the \texttt{\langle property list \rangle} is empty the result is empty. This is not available in older versions of Xe\LaTeX{}.

\textbf{\TeX}Xhackers note: The result is returned within the \texttt{\unexpanded} primitive \texttt{\exp_not:n}, which means that the \texttt{\langle value \rangle} does not expand further when appearing in an x-type argument expansion.
37.11 Additions to \texttt{l3seq}

\begin{itemize}
\item \seqsetthread\texttt{NNN} \star \seqsetthread\texttt{(N|N|c|N|cc|N)} \star
\end{itemize}

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 (i.e. whichever sequence has fewer items determines how many iterations occur).

\begin{itemize}
\item \seqsetfilter\texttt{NNn} \seqgsetfilter\texttt{NNn}
\end{itemize}

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}.

\texttt{TExhackers note:} Contrarily to other mapping functions, \texttt{seq map break}: cannot be used in this function, and would lead to low-level \texttt{TEx} errors.

\begin{itemize}
\item \seqsetfromfunction\texttt{NnN} \seqgsetfromfunction\texttt{NnN}
\end{itemize}

Sets the \texttt{seq var} 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 {final value}}.

\begin{itemize}
\item \seqsetfrominline\texttt{x:NNn} \seqgsetfrominline\texttt{x:NNn}
\end{itemize}

Sets the \texttt{seq var} 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 {final value}}}, but not the analogous “inline” mappings.
Removes the item of \( \langle \text{sequence} \rangle \) at the position given by evaluating the \( \langle \text{integer expression} \rangle \) and replaces it by \( \langle \text{item} \rangle \). Items are indexed from 1 on the left/top of the \( \langle \text{sequence} \rangle \), or from \(-1\) on the right/bottom. If the \( \langle \text{integer expression} \rangle \) is zero or is larger (in absolute value) than the number of items in the sequence, the \( \langle \text{sequence} \rangle \) is not modified. In these cases, \seq_set_item:Nnn raises an error while \seq_set_item:NnnTF runs the \( \langle \text{false code} \rangle \). In cases where the assignment was successful, \( \langle \text{true code} \rangle \) is run afterwards.

Removes the \( \langle \text{item} \rangle \) at position \( \langle \text{integer expression} \rangle \) in the \( \langle \text{sequence} \rangle \), and places it in the \( \langle \text{token list variable} \rangle \). Items are indexed from 1 on the left/top of the \( \langle \text{sequence} \rangle \), or from \(-1\) on the right/bottom. If the position is zero or is larger (in absolute value) than the number of items in the sequence, the \( \langle \text{seq var} \rangle \) is not modified, the \( \langle \text{token list} \rangle \) is set to the special marker \q_no_value, and the \( \langle \text{false code} \rangle \) is left in the input stream; otherwise the \( \langle \text{true code} \rangle \) is. The \( \langle \text{token list} \rangle \) assignment is local while the \( \langle \text{sequence} \rangle \) is assigned locally for \texttt{pop} or globally for \texttt{gpop} functions.

## 37.12 Additions to \texttt{l3sys}

The version string of the current engine, in the same form as given in the banner issued when running a job. For \texttt{pdfTeX} and \texttt{LuaTeX} this is of the form

\[
\langle \text{major} \rangle.\langle \text{minor} \rangle.\langle \text{revision} \rangle
\]

For \texttt{XeTeX}, the form is

\[
\langle \text{major} \rangle.\langle \text{minor} \rangle
\]

For \texttt{pTeX} and \texttt{upTeX}, only releases since \texttt{TeX} Live 2018 make the data available, and the form is more complex, as it comprises the \texttt{pTeX} version, the \texttt{uTeX} version and the \texttt{e-pTeX} version.

\[
p\langle \text{major} \rangle.\langle \text{minor} \rangle.\langle \text{revision} \rangle-u\langle \text{major} \rangle.\langle \text{minor} \rangle-e\langle \text{pTeX} \rangle
\]

where the \texttt{u} part is only present for \texttt{upTeX}.

Tests if the engine has a pseudo-random number generator. Currently this is the case in \texttt{pdfTeX}, \texttt{LuaTeX}, \texttt{pTeX}, \texttt{uTeX} and recent releases of \texttt{XeTeX}.
37.13 Additions to \l3tl

\tl_range_braced:Nnn * \tl_range_braced:cnn * \tl_range_braced:nnn * \tl_range_unbraced:Nnn * \tl_range_unbraced:cnn * \tl_range_unbraced:nnn *

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 \tl_range:nnn. Spaces are ignored. Regardless of whether items appear with or without braces in the \langle token list \rangle, the \tl_range_braced:nnn function wraps each item in braces, while \tl_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 \{b}{c}{d}{e{}}, \{c}{d}{e{}}{f}, \{e{}}{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 bcde{}, cdef, e{}f, and an empty line to the terminal. Because braces are removed, the result of \tl_range_unbraced:nnn may have a different number of items as for \tl_range:nnn or \tl_range_braced:nnn. In cases where preserving spaces is important, consider the slower function \tl_range:nnn.

\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.

\begin{verbatim}
\tl_build_begin:N \tl_build_gbegin:N
\end{verbatim}

Clears the \langle tl var \rangle and sets it up to support other \tl_build_... functions, which allow accumulating large numbers of tokens piece by piece much more efficiently than standard \l3tl functions. Until \tl_build_end:N \langle tl var \rangle is called, applying any function from \l3tl other than \tl_build_... will lead to incorrect results. The begin and gbegin functions must be used for local and global \langle tl var \rangle respectively.

\begin{verbatim}
\tl_build_clear:N \tl_build_gclear:N
\end{verbatim}

Clears the \langle tl var \rangle and sets it up to support other \tl_build_... functions. The clear and gclear functions must be used for local and global \langle tl var \rangle respectively.
\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

\tl_build_put_left:Nn (\tl_var) \{\tokens\}
\tl_build_put_right:Nn (\tl_var) \{\tokens\}

Adds \tokens{} to the left or right side of the current contents of \tl_var. The \tl_var must have been set up with \tl_build_begin:N or \tl_build_gbegin:N. The put and gput functions must be used for local and global \tl_var respectively. The right functions are about twice faster than the left functions.

\tl_build_get:N \tl_build_get:NN

Stores the contents of the \tl_var{} in the \tl_var{}. The \tl_var{} must have been set up with \tl_build_begin:N or \tl_build_gbegin:N. The \tl_var{} is a “normal” token list variable, assigned locally using \tl_set:Nn.

\tl_build_end:N \tl_build_gend:N

\tl_build_end:N (\tl_var)
Gets the contents of \tl_var and stores that into the \tl_var using \tl_set:Nn or \tl_gset:Nn. The \tl_var must have been set up with \tl_build_begin:N or \tl_build_gbegin:N. The end and gend functions must be used for local and global \tl_var respectively. These functions completely remove the setup code that enabled \tl_var to be used for other \tl_build_... functions.

37.14 Additions to \texttt{l3token}

\c_catcode_active_space_tl

Token list containing one character with category code 13, (“active”), and character code 32 (space).

\char_to_utfviii_bytes:n \char_to_utfviii_bytes:n {\codepoint}

Converts the (Unicode) \texttt{\codepoint} 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 \char_to_nfd:N {\char}

Converts the \texttt{\char} to the Unicode Normalization Form Canonical Decomposition. The category code of the generated character is the same as the \texttt{\char}. 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 \textit{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 \textit{inline code} as \texttt{#1}. When begin-group or end-group tokens (usually \texttt{\{ 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{verbatim}
\peek_catcode_collect_inline:Nn A { \iow_term:n {#1} } Hello,~world!
\end{verbatim}

Another example is that the following code tests if the next token is *, ignoring intervening spaces, but putting them back using \texttt{#1} if there is no *.

\begin{verbatim}
\peek_meaning_collect_inline:Nn \c_space_token { \peek_charcode:NTF * { star } { no-star #1 } }
\end{verbatim}

Removes explicit and implicit space tokens (category code 10 and character code 32) from the input stream, then inserts \textit{\texttt{\{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>atan</th>
<th>atand</th>
<th>\bool commands:</th>
</tr>
</thead>
<tbody>
<tr>
<td>!</td>
<td>251</td>
<td>254</td>
<td>\bool_case_true:n</td>
</tr>
<tr>
<td>&amp;h</td>
<td>250</td>
<td>254</td>
<td>\bool_case_true:NTF</td>
</tr>
<tr>
<td>*</td>
<td>251</td>
<td>254</td>
<td>\bool_case_false:n</td>
</tr>
<tr>
<td>**</td>
<td>251</td>
<td>254</td>
<td>\bool_case_false:NTF</td>
</tr>
<tr>
<td>+</td>
<td>250, 251</td>
<td>303</td>
<td>\bool_constant:N</td>
</tr>
<tr>
<td>-</td>
<td>250, 251</td>
<td>303</td>
<td>\bool_constant:N</td>
</tr>
<tr>
<td>/</td>
<td>42</td>
<td>63</td>
<td>\bool_do_until:N</td>
</tr>
<tr>
<td>:N</td>
<td>42</td>
<td>63</td>
<td>\bool_do_until:nn</td>
</tr>
<tr>
<td>:V</td>
<td>42</td>
<td>66</td>
<td>\bool_do_until:nn</td>
</tr>
<tr>
<td>:V_unbraced</td>
<td>42</td>
<td>66</td>
<td>\bool_do_until:nn</td>
</tr>
<tr>
<td>:e</td>
<td>42</td>
<td>67</td>
<td>\bool_do_while:N</td>
</tr>
<tr>
<td>:e_unbraced</td>
<td>42</td>
<td>67</td>
<td>\bool_do_while:nn</td>
</tr>
<tr>
<td>:error</td>
<td>81</td>
<td>63</td>
<td>\bool_gset:N</td>
</tr>
<tr>
<td>:f</td>
<td>42</td>
<td>63</td>
<td>\bool_gset_eq:NN</td>
</tr>
<tr>
<td>:f_unbraced</td>
<td>42</td>
<td>63</td>
<td>\bool_gset_eq:NN</td>
</tr>
<tr>
<td>:n</td>
<td>42</td>
<td>66</td>
<td>\bool_gset_inverse:N</td>
</tr>
<tr>
<td>:o</td>
<td>42</td>
<td>66</td>
<td>\bool_gset_inverse:N</td>
</tr>
<tr>
<td>:o_unbraced</td>
<td>42</td>
<td>66</td>
<td>\bool_gset_inverse:N</td>
</tr>
<tr>
<td>:p</td>
<td>42</td>
<td>63</td>
<td>\bool_if:N</td>
</tr>
<tr>
<td>:v</td>
<td>42</td>
<td>63, 65, 67</td>
<td>\bool_if:nTF</td>
</tr>
<tr>
<td>:v_unbraced</td>
<td>42</td>
<td>63, 65, 67</td>
<td>\bool_if:nTF</td>
</tr>
<tr>
<td>:x</td>
<td>42</td>
<td>64</td>
<td>\bool_if_exist:p:N</td>
</tr>
<tr>
<td>:x_unbraced</td>
<td>42</td>
<td>64</td>
<td>\bool_if_exist:p:N</td>
</tr>
<tr>
<td>&lt;</td>
<td>250</td>
<td>65</td>
<td>\bool_if_p:n</td>
</tr>
<tr>
<td>=</td>
<td>250</td>
<td>65, 66</td>
<td>\bool_if_all:NTF</td>
</tr>
<tr>
<td>&gt;</td>
<td>250</td>
<td>65, 66</td>
<td>\bool_if_all:ntF</td>
</tr>
<tr>
<td>?</td>
<td>250</td>
<td>65, 66</td>
<td>\bool_if_and:NTF</td>
</tr>
<tr>
<td>?:</td>
<td>250</td>
<td>65, 66</td>
<td>\bool_if_and:nF</td>
</tr>
<tr>
<td>^</td>
<td>251</td>
<td>65, 66</td>
<td>\bool_if_or:NTF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>250</td>
</tr>
<tr>
<td>\bool_log:N</td>
<td>251</td>
<td>64</td>
<td>\bool_log:n</td>
</tr>
<tr>
<td>\bool_log:n</td>
<td>253</td>
<td>64</td>
<td>\bool_log:n</td>
</tr>
<tr>
<td>\bool_new:N</td>
<td>253</td>
<td>63</td>
<td>\bool_new:N</td>
</tr>
<tr>
<td>\bool_not:p:n</td>
<td>254</td>
<td>66</td>
<td>\bool_not:p:n</td>
</tr>
<tr>
<td>\bool_set:N</td>
<td>253</td>
<td>63</td>
<td>\bool_set:N</td>
</tr>
<tr>
<td>\bool_set_eq:NN</td>
<td>253</td>
<td>63</td>
<td>\bool_set_eq:NN</td>
</tr>
<tr>
<td>\bool_set_inverse:N</td>
<td>253</td>
<td>63</td>
<td>\bool_set_inverse:N</td>
</tr>
<tr>
<td>\bool_set_inverse:N</td>
<td>253</td>
<td>302</td>
<td>\bool_set_inverse:N</td>
</tr>
<tr>
<td>\bool_set_true:N</td>
<td>253</td>
<td>63</td>
<td>\bool_set_true:N</td>
</tr>
</tbody>
</table>

309
\dim_set:Nn ........................................ 207  \dim_set_eq:NN .................................... 207  \dim_set:NN ......................................... 207  \dim_sign:n ......................................... 212  \dim_step:nn ....................................... 210  \dim_step_function:nnnN ......................... 211  \dim_step_variable:nnnN ......................... 212  \dim_sub:N .......................................... 207  \dim_to_decimal:n ................................ 212  \dim_to_decimal_in_sp:n ......................... 210  \dim_to_decimal_in_bp:n ......................... 210  \dim_zero_new:N ................................... 206  \dim_zero:N ......................................... 206  \dim_while:n ...................................... 211  \\

draw commands:  \\
\draw_begin: ...................................... 290  \draw_end: ......................................... 290

E

else commands:  \\
\else: 27, 62, 68, 91, 166, 167, 220, 280
\else: ............................................. 256  \else: ............................................. 256  \else: ............................................. 251

exp commands:  \\
\exp:w ............................................ 41, 42  \exp:w ............................................ 38, 40, 41, 193  \exp:w ............................................ 34  \exp:w ............................................ 32, 34  \exp:w ............................................ 36  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 36  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 36  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \exp:w ............................................ 37  \\
\exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  \exp_args:NN ................. 30, 36  
\prop_pop:NnNTF 200, 202
\prop_put:Nn 226
\prop_put:Nn 199, 200
\prop_put_from_keyval:Nn 200
\prop_put_if_new:Nn 199
\prop_rand_key_value:N 303
\prop_remove:Nn 201
\prop_set_eq:NN 199
\prop_set_from_keyval:Nn 199, 200
\prop_show:N 204
\g_tampa_prop 204
\l_tampa_prop 204
\l_tamp_prop 204
\ProvidesExplClass 9
\ProvidesExplFile 9
\ProvidesExplPackage 9
pt 256

Q

\quark commands:
\q_mark 136
\q_nil 25, 116, 136
\q_no_value ... 72, 85, 86, 92–94,
135, 136, 142, 143, 149, 177, 200, 305
\quark_if_nil:NTF 136
\quark_if_nil:nTF 136
\quark_if_nil:p:N 136
\quark_if_nil:p:n 136
\quark_if_no_value:NTF 136
\quark_if_no_value:nTF 136
\quark_if_no_value:p:N 136
\quark_if_no_value:p:n 136
\quark_if_recursion_tail_-break:NN 138
\quark_if_recursion_tail_-break:nN 138
\quark_if_recursion_tail_stop:N 137
\quark_if_recursion_tail_stop:n 137
\quark_if_recursion_tail_stop_-do:Nn 137
\quark_if_recursion_tail_stop_-do:n 137
\quark_new:N 136
\q_recursion_stop 25, 137, 138
\q_recursion_tail 137, 138
\q_stop ... 25, 38, 112, 135, 136

R

\rand 255
\randint 255

\regex commands:
\regex_const:Nn 46, 54
\regex_count:Nn 55

\seq commands:
\c_empty_seq 151
\seq_clear:N 140, 151
\seq_clear_new:N 140
\seq_concat:NNN 141, 151
\seq_const_from_clist:N 141
\seq_count:N 143, 148, 150, 235
\seq_gclear:N 140
\use:n \ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\.