Prototype reimplementation of $\LaTeX 2_{\varepsilon}$'s block environments using templates

ĿÆT_EX Project*

v0.9j 2025-07-09

Abstract

Contents

9.1	Template types and templates for blocks and lists 2.1 Template types						
_	2.1.1	The template type 'block'					
_	2.1.2	The template type 'para'					
	2.1.3	The template type 'list'					
_	2.1.4	The template type 'item'					
	2.1.5	The template type 'blockeny'					
		$_{ m iplates}$					
2	2.2.1	The blockenv template 'display'					
2	2.2.2	The block template 'display'					
2	2.2.3	The para template 'std'					
2	2.2.4	The list template 'std'					
2	2.2.5	The item template 'std'					
Dec	Declaration of standard block environments						
3.1	The	center, flushleft, and flushright environments					
2.0	The	quote and quotation environments					
3.2	The	verbatim and verbatim* environments					
3.3	1110	VCIDACIM AND VCIDACIM CHVITOHIHCHOS					
_		itemize environment					
3.3	The	itemize environment					
3.3 3.4	The The	itemize environment					
3.3 3.4 3.5	The The	<pre>enumerate environment</pre>					
3.3 3.4 3.5 3.6	The The The	<pre>itemize environment</pre>					
3.3 3.4 3.5 3.6 3.7	The The The The	<pre>enumerate environment</pre>					

^{*}Initial reimplementation of lists done by Bruno Le Floch, generalized second version with tagging support by Frank Mittelbach.

5	Tagging support5.1 Paragraph tags5.2 Tagging recipes					
3	Debugging					
7	New and redefined kernel command					
Ind	lex	15				

1 Introduction

The list implementation in \LaTeX $2_{\mathcal{E}}$ serves a dual purpose: it implements real lists such as itemize or enumerate, but it is also used as the basis for vertical blocks, i.e., to specify the vertical spacing and paragraph handling after such block, e.g., in environments like center, quote, verbatim, or in the theorem environments. They are all implemented as "trivial" lists with a single (hidden) item.

While this was convenient to get a consistent layout using a single implementation it is not adequate if it comes to interpreting the structure of a document, because environments based on trivlist should not advertise themselves as being a "list" — after all, from a semantic point of view they aren't lists.

The approach taking here is therefore to offer separate template types: block (horizontally or vertically oriented data that needs some handling at the start and the end), para (that deals with different paragraph layouts), list (that handles list related parameters, and item (for item layouts and handling). To address the independent aspects we have the template type blockenv that ties them together as necessary.

For example, a quote environment would make use of a (display) block and some para instance while an standard enumerate would make use of a display block, a list, and an item and para instance. An inline list (like enumerate* from the enumitem package) would be using the same list instance but a different (horizontally oriented) block instance build from a different template.

2 Template types and templates for blocks and lists

2.1 Template types

2.1.1 The template type 'block'

Arg: 1 key/value list to alter the default block parameters

Semantics:

Handle the layout aspects of a block of data. In case of a "display" block (i.e., vertically oriented) the spacing and page breaking as well as the handling if the block starts a paragraph or ends one, that is, if text is immediately following the block without being separated by an empty line, then this text is considered to be in the same paragraph as the block.

In case of a horizontally oriented block it covers any special handling at the start and end of the block, e.g, extra spacing, prohibitying or encuraging line breaks, and so forth.

2.1.2 The template type 'para'

Arg: 1 key/value list to alter the default item parameters

Semantics:

Sets up paragraph-specific parameters for H&J, e.g., to implement justification variations, the behavior of \\ etc. The instances are used in higher-level templates, e.g., in a block.

2.1.3 The template type 'list'

Arg: 1 key/value list to alter the default item parameters

Semantics:

Handle the aspects related to list design, e.g., the use and formatting of counters, etc.

Note that this does not cover block-related aspects, i.e., a list instance could be used both for a display list or for an inline line.

2.1.4 The template type 'item'

Arg: 1 key/value list to alter the default item parameters

Semantics:

A sub-type used as part of list to easily cover alternative layout for list items.

2.1.5 The template type 'blockenv'

Arg: 1 key/value list to alter the default parameters of the template instances used by the particular block environment

Semantics:

This template type is used to implement document-level environments. It defines a block instance to handle the layout at the "edge" of the environment data, possibly some paragraph setup through a para instance, potentially an "inner" instance for more complicated environments (such as lists), and possibly some additional setup code for certain environments.

It also defines how the blockenv behaves with respect to nesting, e.g., does it change when nested and if so how many levels of nesting are supported, etc.

Finally, the template type defines how it appears in a tagged PDF document, what tag names are used, how they are rolemapped and whether it adds additional attributes, etc.

2.2 Templates

2.2.1 The blockenv template 'display'

Attributes:

name (tokenlist) Name of the environment used in tracing and error messages.

tag-name (tokenlist) Name of the tag used for the block inside the PDF. If not explicitly given the name is defined by the tagging-recipe. Note that in case of tagging-recipe=basic no tag for the block is produced, so any key settings are ignored.

Default: \(\lambda empty \rangle \)

tag-attr-class (tokenlist) An explicit tag class attribute. Default: (empty)

tagging-recipe (tokenlist) Defines the way tagging is done. Currently the values basic, standard, and list are supported.

Default: standard

increment-level (boolean) Does this blockenv increase the block level if it is nested in an outer block?

Default: true

setup-code (tokenlist) Initial setup code. This is executed after legacy defaults (from \@listi, \@listii, etc.) are used but before the block instance is called.

Default: (empty)

block-instance (tokenlist) Part of the name of the block instance that is called. The full name has a -\left(level\right) appended.

Default: displayblock

para-instance (tokenlist) Paragraph settings to use within the environment. If $\langle empty \rangle$ then outer values are retained. However, the block template resets some values, which may not be the right thing to do.

Default: $\langle empty \rangle$

inner-level-counter (tokenlist) Name of an existing (!) counter that is incremented and used to determine final name of the inner-instance or empty if always the same inner instance should be used.

max-inner-levels (tokenlist) Maximum number of nested environments of this kind.

Only relevant if there is a inner-level-counter specified.

Default: 4

inner-instance-type (tokenlist) Template type of the inner instance. Default: list

inner-instance (tokenlist) Name of the inner instance (if any). If there is an inner-level-counter then the instance name gets -\((counter value \) appended.

Default: (empty)

Default: false

tagging-suppress-paras (boolean) describe

final-code (tokenlist) Final setup code

Default: \ignorespaces

Semantics & Comments: This blockenv template supports the legacy list setting that are found in many document classes in the macros $\$ 0listi, $\$ 0listii, up to $\$ 0listvi. It also uses the counter $\$ 0listdepth to track nesting of block, again mainly to support legacy setups (internally it gives it a more appropriate name but it remains accessible through the $\$ 1FX $\$ 2 $\$ 6 name).

It first checks that nothing is too deeply nested. If the level should increase then the increments the \@listdepth counter and calls the corresponding \@list... macro to update the legacy defaults. If increment-level is set to false this is bypassed.

It then sets up the tagging via the tagging-recipe setting and executes any code in setup-code.

Afterwards it calls the appropriate block instance based on block-instance and current level, e.g., displayblock-1. Then it sets up paragraph parameters if a parainstance was specified (otherwise they stay as they are).

If a inner-instance was specified this is called next, or more precisely: if no inner-level-counter was specified the instance inner-instance is called.

Otherwise, the inner-level-counter is incremented and the instance with the name inner-instance-inner-level-counter is called.

Finally, the final-code is executed (by default \ignorespaces).

The maximum number of blockenvs that can be nested into each other is restricted by the LATEX counter maxblocklevels with a default value of 6. If this value is increased then it is necessary to provide additional instances, e.g., displayblock-7, etc. Decreasing is, of course, always possible, then some of the instances defined are not used and instead the user gets an error that there is too much nesting going on.

If the key increment-level is set to false then such an environment doesn't alter the nesting level and therefore you can nest those environments as often as you like (a typical example would be flushleft anywhere in the nesting hierarchy, that would have no effect on hitting the boundary).

2.2.2 The block template 'display'

Attributes:

```
begin-vspace (skip)
                                                                   Default: \topsep
begin-extra-vspace (skip)
                                                                Default: \partopsep
para-vspace (skip)
                                                                   Default: \parsep
                                                   Default: value from begin-vspace
end-vspace (skip)
end-extra-vspace (skip)
                                            Default: value from begin-extra-vspace
item-vspace (skip) The space in front of an item if the block is a list; if not the setting
     has no effect
                                                                  Default: \itemsep
begin-penalty (integer)
                                                        Default: \@beginparpenalty
end-penalty (integer)
                                                          Default: \@endparpenalty
left-margin (length)
                                                               Default: \leftmargin
right-margin (length)
                                                             Default: \rightmargin
para-indent (length)
                                                                        Default: Opt
```

Semantics & Comments: The idea of a heading key needs some further thoughts and therfore has been removed for now. Maybe instead the template type should accept a second argument and receive input for such a heading from the document level instead.

The names of the keys need further thoughts and some decision. Right now it is a mixture of those with hyphens and those that match legacy register names (the way enumitem did its keys).

2.2.3 The para template 'std'

Attributes:

para-indent (length) Default: \parindent begin-hspace (skip)Default: Opt left-hspace (skip) Default: Opt right-hspace (skip)Default: Opt end-hspace (skip) Default: \Oflushglue Default: false fixed-word-spaces (boolean) Default: 5000 final-hyphen-demerits (integer) newline-cmd (tokenlist) Default: \@normalcr para-attr-class (tokenlist) Default: justify

2.2.4 The list template 'std'

Attributes:

counter (tokenlist) Counter name to be used in a numbered list or empty, if the list is unnumbered

item-label (tokenlist) Label "string" for a fixed label or as generated from the current counter value

start (integer) Start value for the counter if the list is numbered, otherwise irrelevant

Default: 1

resume (boolean) Should a numbered list be resumed from the last instance?

Default: false

item-instance (instance) Instance of type item to be used to format the label string

Default: basic

item-vspace (skip) The space in front of an item in the list. If not specified the value specified in the block template instance is used

item-indent (length) Horizontal displacement of the item. Default: Opt

item-penalty (integer) Penalty for breaking before an item (except the first)

Default: \@itempenalty

label-width (length) Width reserved for the formatted item labelDefault: \labelwidth

label-sep (length) Horizontal separation between label and following text

Default: \labelsep

legacy-support (boolean) Is formatting the label via \makelabel supported?

Default: false

2.2.5 The item template 'std'

Attributes:

counter-label (function1) unused Default: \arabic{#1}

counter-ref (function1) unused Default: value from counter-label

label-ref (function1) unused Default: #1

label-autoref (function1) unused Default: item #1

label-strut (boolean) Add a \strut to the label? Default: false

label-boxed (boolean) Should the label be boxed? Default: true

next-line (boolean) Default: false

text-font (tokenlist) unused

compatibility (boolean) Default: true

Semantics & Comments: This template is only rudimentary implemented at the moment. It probably needs other keys and the existing ones need a proper implementation.

3 Declaration of standard block environments

3.1 The center, flushleft, and flushright environments

The center environment is defined through the blockenv instance center which makes use of the block instance displayblock-\(\lambda level\) and the para instance center. The block nesting level is not incremented. With respect to tagging, text separated by \par commands (or empty lines) inside the environment is not tagged as separate paragraphs, i.e., the whole environment is considered to be part of an outer paragraph. The default implementation is

The flushleft and flushright environments are defined in a similar way.

3.2 The quote and quotation environments

The quote environment is defined through the blockenv instance quote which makes use of the block instance quoteblock- $\langle level \rangle$. The paragraph setup is inherited. The block nesting level is incremented. The default implementation is

The implementation of quotation is similar but uses quotationblock-(level).

3.3 The verbatim and verbatim* environments

Both the verbatim environment is defined through the blockenv instance verbatim which makes use of the block instance verbatimblock- $\langle level \rangle$ and the para instance justify. The block nesting level is not incremented. Verbatim processing requires various catcode changes, etc. and as a consequence a special parsing routine that grabs the whole environment while these catcodes are in force. This setup is done in the final-code key and its last action is to initiate the special parsing. The default implementation is

```
\DeclareInstance{blockenv}{verbatim}{display}
  name
                         = verbatim,
  tag-name
                         = verbatim,
  tagging-recipe
                         = standard,
  tagging-suppress-paras = true,
  increment-level
                         = false,
  block-instance
                         = verbatimblock ,
  para-instance
                         = justify ,
  final-code
                         = \legacyverbatimsetup
                            \@setupverbinvisiblespace \@vobeyspaces
                           \@xverbatim
}
```

The implementation of verbatim* is similar using the blockenv instance verbatim*. Its final-code sets up visible spaces and a slightly different parsing that grabs everything up to \end{verbatim*}. Otherwise the setup is identical.

3.4 The itemize environment

The itemize environment is defined through the blockenv instance itemize which makes use of the block instance list- $\langle level \rangle$, and an inner instance itemize- $\langle inner-level \rangle$ of type list. The paragraph setup is inherited. The $\langle inner-level \rangle$ is determined through $\langle inner-level \rangle$ are incremented. The default implementation is

```
\DeclareInstance{blockenv}{itemize}{display}
{
  name
                       = itemize,
  tag-name
                       = itemize,
  tag-attr-class
                       = itemize,
  tagging-recipe
                       = list,
  inner-level-counter = \@itemdepth,
  increment-level
                       = true,
  max-inner-levels
                       = 4,
  block-instance
                       = listblock ,
  inner-instance
                       = itemize ,
}
```

3.5 The enumerate environment

The enumerate environment is similar to itemize but uses the blockenv instance enumerate, the block instance list- $\langle level \rangle$, and the inner instance enumerate- $\langle inner-level \rangle$. The $\langle inner-level \rangle$ is determined through $\langle enumdepth \rangle$.

3.6 The description environment

The description environment uses the blockenv instance description, the block instance list-\langle level \rangle, and the inner instance description (no dependency on the nesting level), i.e., the environment has the same appearance on all nesting levels.

fix

3.7 The list environment

The generic list environment of \LaTeX 2ε is modeled with a blockenv instance named list, a block instance named list- $\langle level \rangle$, and an inner instance named legacy (with no dependency on the nesting level). This environment has two arguments and customization of the layout is expected to be directly set in the second argument. For this reason this legacy instance is something that shouldn't be changed (all that is attempted to provide a way to support legacy setups).

To set up the default settings (as they were used in \LaTeX 2_{ε}) the setup-code key gets \legacylistsetupcode assigned, so the default setup (that should probably not be changed) looks as follows:

```
\DeclareInstance{blockenv}{list}{display}
{
                       = list,
  name
  tag-name
                       = list,
  tagging-recipe
                       = list,
  increment-level
                       = true,
                       = \legacylistsetupcode ,
  setup-code
  block-instance
                       = listblock ,
  inner-instance
                       = legacy ,
}
```

3.8 The verse environment

The verse environment is currently still implemented as a list without real items (as in $A^{T} \to X 2_{\varepsilon}$. That needs updating.

3.9 The trivlist environment

In \LaTeX 2ε trivlist was used to define various display environments that aren't really lists at all. To support such legacy definitions (even though they shuld be updated to achieve proper tagging) we continue to support and implement it as a list environment with a few hardwired settings mimicking the original behavior.

3.10 Environments declared through \newtheorem

to document

4 Adjusting the layout of standard block environments

to document

5 Tagging support

5.1 Paragraph tags

Paragraphs in IATEX can be nested, e.g., you can have a paragraph containing a display quote, which in turn consists of more than one (sub)paragraph, followed by some more text which all belongs to the same outer paragraph.

In the PDF model and in the HTML model that is not supported — a limitation that conflicts with real live, given that such constructs are quite normal in spoken and written language.

The approach we take to resolve this is to model such "big" paragraphs with a structure named <text-unit> and use <text> (rollmapped to <P>) only for (portions of) the actual paragraph text in a way that the <text>s are not nested. As a result we have for a simple paragraph the structures

```
<text-unit>
  <text>
    The paragraph text ...
  </text>
  </text-unit>
```

The <text-unit> structure is rollmapped to <Part> or possibly to <Div> so we get a valid PDF, but processors who care can identify the complete paragraphs by looking for <text-unit> tags.

In the case of an element, such as a display quote or a display list inside the paragraph, we then have

```
<text-unit>
    <text>
        The paragraph text before the display element ...
    </text>
        <display element structure>
            Content of the display structure possibly involving inner <text-unit> tags
        </display element structure>
        <text>
            ... continuing the outer paragraph text
        </text>
    </text-unit>
```

In other words such a display block is always embedded in a <text-unit> structure, possibly preceded by a <text>...</text> block and possibly followed by one, though both such blocks are optional.

Thus an **itemize** environment that has some introductory text but no text immediately following the list would be tagged as follows:

```
<text-unit>
  <text>
    The intro text for the itemize environment ...
  </text>
```

The <itemize> is rollmapped to <L>.

For some display blocks, such as centered text, we use a simpler strategy. Such blocks still ensure that they are inside a <text-unit> structure but their body uses simple <text> blocks and not <text-unit> ctext> inside, e.g., the input

```
This is a paragraph with some
\begin{center}
   centered lines
   with a paragraph break between them
\end{center}
followed by some more text.
will be tagged as follows:
  <text-unit>
    <text>
      This is a paragraph with some
    </text>
    <text /0 /Layout /TextAlign/Center>
      centered lines
    </text>
    <text /0 /Layout /TextAlign/Center>
      with a paragraph break between them
    </text>
    <text>
      followed by some more text.
  </text-unit>
```

5.2 Tagging recipes

There are a number of different tagging recipes that implement different tagging approaches. They are selected through the tagging-recipe of the blockenv template. Currently the following values are implemented:

standalone This recipe does the following:

- Ensure that the blockenv is not inside a <text-unit> structure. If necessary, close the open one (and any open <text> structure).
- Text inside the body of the environment start with <text-unit><text> unless the key tagging-suppress-paras is set to true (which is most likely the wrong thing to do because we then get just <text> as the structure).
- At the end of the environment close </text> and possibly an inner </text-unit> if open.
- Finally, ensure that after the environment a new <text-unit> is started, if appropriate, e.g., if text is following.

basic This recipe does the following:

- Ensure that the blockenv is inside a <text-unit> structure, if necessary, start one.
- If inside a <text-unit><text>, then close the </text> but leave the <text-unit> open.
- Text inside the body of the environment start with <text-unit><text> if tagging-suppress-paras is set to false, otherwise just with <text>.
- At the end of the environment close </text> and possibly an inner </text-unit> if open.
- Then look if the environment is followed by an empty line (\par). If so, close the outer </text-unit> and start any following text with <text-unit><text>. Otherwise, don't and following text restarts with a just a <text> (and no paragraph indentation)

standard This recipe is like the basic one as far as handling <text-unit> and <text> is concerned. In addition

- it starts an inner tagging structure (i.e., which is therefore a child of the outer <text-unit>).
- By default this structure is a <Div> unless overwritten by the key tag-name. If that key is used, a suitable rolemap needs to be provided for the name given.
- At the end of the environment that inner structure is closed again so that we are back on the <text-unit> level from the outside.
- Then the lookahead for an empty line is done as described previously.

list This recipe is like the standard one except that

- the inner structure is a list (<L>).
- Furthermore everything is set up so that we have list items () with suitable substructures (<Lbl> for the item labels and <LBody> for the item bodies).
- If the key tag-name is specified, this is used as the tag name for the whole list instead of <L>. Of course, it should then have a suitable rollmap.
- If the key tag-attr-class is specified then this is used as the class attribute. Again, this requires a suitable setup on the outside.

- At the end of the environment the $\$, $\$, and $\$ (or the tag name used) are closed.
- Then the lookahead for an empty line is done as described previously.

6 Debugging

\DebugBlocksOn
\DebugBlocksOff
\block_debug_on:
\block_debug_off

These commands enable/disable debugging messages.

	7 New and redefined kernel command
\@doendpe	The original IATEX $2_{\mathcal{E}}$ command is augmented to allow for tagging.
\legacyverbatimsetup \legacylistsetupcode	to be documented
\@setupverbinvisiblespace	A counterpart definition to the kernel command \@setupverbinvisiblespace, needed as we need to handle real space chars in verbatim.
endblockenv ag_block_nesting_depth_int	to be documented
\newtheorem \@thm \@begintheorem	Redefined to make theorems tagging aware.
\item \@itemlabel	The \item is redefined.
\c@maxblocklevels	A counter to increase or decrease the number of supported level. If increased, one needs to supply additional level instances.

 $\verb|\begin| The \verb|\begin| is slightly redefined to handle \verb|\alpha| better. TODO: move to kernel |$

\para_end: TODO: consider name, document

para/begin The para/begin hook is enhanced to support list ends

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.

В	\@begintheorem
\begin 14	\@doendpe
block commands:	\@endparpenalty 5
\block_debug_off: 14	\@enumdepth g
\block_debug_on: 14	\@flushglue 6
\g_block_nesting_depth_int 14	\@itemdepth 9
_	\@itemlabel 14
D	\@itempenalty 7
\DebugBlocksOff 14	\@list 5
\DebugBlocksOn 14	\@listdepth 5
	\@listi
${f E}$	\@listii
endblockenv	\@listvi 5
	\@normalcr 6
I	\@setupverbinvisiblespace 14
\item	\@thm 14
	\arabic
${f L}$	\begin 14
\legacylistsetupcode 10, 14	\c@maxblocklevels 14
\legacyverbatimsetup 14	\ignorespaces 4, 5
3. 7	\item 14
N	\itemsep 5
\newtheorem 10, 14	\labelsep
Th.	\labelwidth 7
P	\leftmargin 5
\par 8	\makelabel
para commands:	\par 13
\para_end: 15	\parindent 6
para/begin	\parsep 5
	\partopsep 5
T	\rightmargin 5
T _E X and L ^{\perp} T _E X 2_{ε} commands:	\strut 7
\c 0beginparpenalty 5	\topsep 5