Prototype reimplementation of LaTeX 2_ε's role mapping

LaTeX Project, initial implementation Ulrike Fischer

v0.8r 2026-09-27

1 Introduction

This file provides support for tag names and symbolic structure names.

At first, it creates tagpdf-ns-latex-lab.def which collects tag names from the LaTeX namespace and their role mapping. The names and the roles are not necessarily the final ones!

Later settings of the same tag overwrite previous settings. The syntax of a line is a list of three strings ending with commas:

```
latex-tag, PDF-tag, name space,
```

The name space is normally pdf2 with the exception of the following PDF-tags which are only in pdf namespace: Art, BlockQuote, TOC, TOCI, Index, Private, Quote, Note, Reference, BibEntry, Code.

At second it provides an interface to use symbolic structure names which can be used in the `tag` key of `\tagstructbegin`. The symbolic names are strings structured with slashes similar to the names of (tagging) sockets. The list of such names used by LaTeX is below in the implementation section.

In PDF 2.0 there are namespaces, so theoretically it is possible to use the same tag name for different purposes and with different role mapping. But practically this is not an option as we build also a global RoleMap dictionary for compatibility with applications that are not PDF 2.0 aware, and this dictionary can only contain unique tag names. So we assume that tag names are used only once across all namespaces. Documents that want to use their own specialized tag names must take care not to clash with existing names, e.g., by using a specific prefix on all their tag names.

`\NewStructureName` `\NewStructureName{`⟨*symbolic name*⟩`}`

This declares a new symbolic structure name ⟨*symbolic name*⟩ which is initially defined as `NonStruct`. To assign it another role use `\AssignStructureRole` described below.

`\UseStructureName` ⋆ `\UseStructureName{`⟨*symbolic name*⟩`}`

This uses the new symbolic structure name ⟨*symbolic name*⟩, e.g. `tagstructbegin{tag=\UseStructureName{sec/section}}` is the same as `tagstructbegin{tag=Sect}` assuming that `\NewStructureName{sec/section}` and `\AssignStructureRole{sec/section}{Sect}` has been used before.

\AssignStructureRole

\AssignStructureRole{⟨*symbolic name*⟩}{⟨*role*⟩}

This assigns to the symbolic structure name ⟨*symbolic name*⟩ the role ⟨*role*⟩. ⟨*symbolic name*⟩ must have been declared first with \NewStructureName. If ⟨*role*⟩ is a known role it is simply used. If it is unknown, it is declared first and rolemapped to the tag currently stored for ⟨*name*⟩. So

```
\NewStructureName{sec/section}
% structure with tag 'NonStruct'
\tagstructbegin{tag=\UseStructureName{sec/section}}

\AssignStructureRole{sec/section}{Sect}
% structure with tag 'Sect'
\tagstructbegin{tag=\UseStructureName{sec/section}}

\AssignStructureRole{sec/section}{Abschnitt}
% structure with tag 'Abschnitt' rolemapped to 'Sect'
\tagstructbegin{tag=\UseStructureName{sec/section}}

\AssignStructureRole{sec/section}{Figure}
% structure with tag 'Figure'
\tagstructbegin{tag=\UseStructureName{sec/section}}
```

Assigning a symbolic structure name a role that is not suitable for the structure element can lead to invalid tagging and parent-child warnings!

```
⟨*ns-latex-lab⟩
%% \ProvidesExplFile {tagpdf-ns-latex-lab.def} {2023-09-04} {0.8}
%% {latex} {https://www.latex-project.org/ns/lab/2022}{}
```

1.1 Roles in the LaTeX name space

```
title-block,  Div,    pdf2,
title-text,   Title,  pdf2,
title-author, P,  pdf2,
title-date,   P,  pdf2,
part,         Part,     pdf2,
part-heading, Title,    pdf2,
```

Headings differ in article and book class. This here is for article:

```
section,      Sect,       pdf2,
subsection,   Sect,       pdf2,
subsubsection,Sect,       pdf2,
paragraph,    Sect,       pdf2,
subparagraph, Sect,       pdf2,
section-heading,      H1,       pdf2,
subsection-heading,   H2,       pdf2,
subsubsection-heading,H3,       pdf2,
paragraph-heading,    H4,       pdf2,
subparagraph-heading, H5,       pdf2,
```

```
list,         L,        pdf2,
itemize,      L,        pdf2,
enumerate,    L,        pdf2,
```

```
description,  L,        pdf2,
quote,        BlockQuote,pdf,
quotation,    BlockQuote,pdf,
item,         LI,       pdf2,
itemlabel,    Lbl,      pdf2,
itembody,     LBody,    pdf2,
footnote,     FENote,   pdf2,
footnotemark, Lbl,      pdf2,
footnotelabel,Lbl,      pdf2,
semantic-para,Part,     pdf2,
text-block,   P,        pdf2,
text-fragment,Span,     pdf2,

text-parbox,     Div,   pdf2,
text-minipage,   Div,   pdf2,
heading-parbox,  Span,  pdf2,
heading-minipage,Span,  pdf2,
```

1.2 From sec code

don't think this should be called `section-number`. Maybe `heading-number`?

```
heading-number, Lbl,    pdf2,
heading-fragment, Span, pdf2,
```

1.3 From block code

```
theorem-like, Sect,     pdf2,
verbatim,     Code,     pdf,
codeline,     Sub,      pdf2,
```

1.4 From float code

```
float,        Aside,    pdf2,
figures,      Sect,     pdf2,
tables,       Sect,     pdf2,

⟨/ns-latex-lab⟩
```

This are the headings used in book

```
⟨*ns-latex-lab-book⟩
%% \ProvidesExplFile {tagpdf-ns-latex-lab-book.def} {2026-08-21} {1.0e}
%% {latex-book} {https://www.latex-project.org/ns/book}{}
chapter,        Sect,pdf2,
chapter-heading,       H1,pdf2,
section-heading,       H2,pdf2,
subsection-heading,    H3,pdf2,
subsubsection-heading, H4,pdf2,
paragraph-heading,     H5,pdf2,
subparagraph-heading,  H6,pdf2,

⟨/ns-latex-lab-book⟩

⟨*package⟩
\ProvidesExplPackage {latex-lab-testphase-names} {\ltlabnamesdate} {\ltlabnamesversion}
  {Code related to the names of tags}
```

```
⟨@@=tag⟩
```

\AssignStructureRole

We redefine \AssignStructureRole to set up roles for tagging. The second argument needs expansion so that we do not store things like \UseStructureName{block/verbatim/codeline}} but its final value. Otherwise …

fix implementation and docu

```
\tl_new:N\l__tag_namespace_tag_tl
\tl_new:N\l__tag_namespace_NS_tl
\tl_new:N\l__tag_namespace_roletag_tl
\tl_new:N\l__tag_namespace_roleNS_tl

\cs_set_protected:Npn\AssignStructureRole#1#2 {
  \exp_args:Nne \__tag_AssignStructureRole {#1}{#2}
}
```

This assigns the target tag name to the symbolic name. The assignment itself is short. The main part of the code is to check that the target actually exists and if not to create a suitable default tag with a suitable role mapping. The tag name can have the form tag/NS. The following table shows how the name space and role mapping of the new tag are then defined:

value of symbolic name	Argument	NS of new tag	role mapping of newtag
oldtag	newtag	user	oldtag/defaultNS
oldtag/oldNS	newtag	user	oldtag/oldNS
oldtag	newtag/newNS	newNS	oldtag/defaultNS
oldtag/oldNS	newtag/newNS	newNS	oldtag/oldNS
oldtag	newtag/unknownNS	user	oldtag/defaultNS
oldtag/oldNS	newtag/unknownNS	user	oldtag/oldNS

```
\cs_set_protected:Npn\__tag_AssignStructureRole#1#2
 {
   \tag_set_split:NNe\l__tag_namespace_tag_tl\l__tag_namespace_NS_tl{#2}
```

this test covers both empty and unknown namespaces:

```
   \tag_if_NS_known:nTF { \l__tag_namespace_NS_tl }
    {
      \tag_if_tag_known:eeF {\l__tag_namespace_tag_tl}{\l__tag_namespace_NS_tl}
       {
         \tag_set_split:NNe
           \l__tag_namespace_roletag_tl
           \l__tag_namespace_roleNS_tl
           {\UseStructureName{#1}}
         \tl_if_empty:NTF \l__tag_namespace_roleNS_tl
          {
            \exp_args:Ne
            \tagpdfsetup
              {
                role/new-tag=
                 {
                   tag = \l__tag_namespace_tag_tl,
                   tag-namespace=\l__tag_namespace_NS_tl,
                   role= \l__tag_namespace_roletag_tl,
                 }
              }
          }
          {
```

```
              \exp_args:Ne
              \tagpdfsetup
                {
                  role/new-tag=
                   {
                     tag = \l__tag_namespace_tag_tl,
                     tag-namespace=\l__tag_namespace_NS_tl,
                     role= \l__tag_namespace_roletag_tl,
                     role-namespace= \l__tag_namespace_roleNS_tl
                   }
                }
          }
      }
    }
    {
      \tag_if_tag_known:eF {\l__tag_namespace_tag_tl}
        {
         \tag_set_split:NNe
           \l__tag_namespace_roletag_tl
           \l__tag_namespace_roleNS_tl
           {\UseStructureName{#1}}
         \tl_if_empty:NTF \l__tag_namespace_roleNS_tl
          {
            \exp_args:Ne
             \tagpdfsetup{role/new-tag=\l__tag_namespace_tag_tl/\l__tag_namespace_roletag_tl}
          }
          {
            \exp_args:Ne
             \tagpdfsetup
              {
               role/new-tag=
                 {
                   tag = \l__tag_namespace_tag_tl,
                   role= \l__tag_namespace_roletag_tl,
                   role-namespace=\l__tag_namespace_roleNS_tl
                 }
              }
          }
      }
    }
  \tl_set:cn { l__tag_name_#1_tl }{#2}
}
```

(*End of definition for* `\AssignStructureRole`*. This function is documented on page 2.*)

⟨/package⟩

1.5 Symbolic names

These are the planned symbolic structure names. Every commented name is not decided and not used yet and can change in future.

1.5.1 Generic

In various places we have to use a Span structure to add attributes.

```
⟨*package⟩
\NewStructureName{span}                % <-- different name?
%\AssignStructureRole{span}{Span}
\AssignStructureRole{span}{text-fragment}
⟨/package⟩
```

1.5.2 Paragraphs

```
⟨*package⟩
\NewStructureName{para/textblock}
\AssignStructureRole{para/textblock}{text-block}
\NewStructureName{para/semantic}
\AssignStructureRole{para/semantic}{semantic-para}
⟨/package⟩
```

1.5.3 Sectioning commands

```
⟨*package⟩
%-1=part, 0=chapter, 1=section etc
\NewStructureName{sec/-1}
\AssignStructureRole{sec/-1}{part}
\NewStructureName{sec/-1/title}
\AssignStructureRole{sec/-1/title}{part-heading}
\NewStructureName{sec/-1/fragment}
```

Sub does not work as part is rolemapped to Title so we use Span

```
\AssignStructureRole{sec/-1/fragment}{heading-fragment}

\NewStructureName{sec/-1/number}
\AssignStructureRole{sec/-1/number}{heading-number}
\NewStructureName{sec/0}
\AssignStructureRole{sec/0}{chapter}
\NewStructureName{sec/0/title}
\AssignStructureRole{sec/0/title}{chapter-heading}
\NewStructureName{sec/0/fragment}
\AssignStructureRole{sec/0/fragment}{heading-fragment}
\NewStructureName{sec/0/number}
\AssignStructureRole{sec/0/number}{heading-number}
\NewStructureName{sec/1}
\AssignStructureRole{sec/1}{section}
\NewStructureName{sec/1/title}
\AssignStructureRole{sec/1/title}{section-heading}
\NewStructureName{sec/1/fragment}
\AssignStructureRole{sec/1/fragment}{heading-fragment}
\NewStructureName{sec/2/fragment}
\NewStructureName{sec/1/number}
\AssignStructureRole{sec/1/number}{heading-number}
\NewStructureName{sec/2}
\AssignStructureRole{sec/2}{subsection}
\NewStructureName{sec/2/title}
\AssignStructureRole{sec/2/title}{subsection-heading}
```

2026/08/22: I have some serious doubts that defining the structure name for the /number on a heading level is in any way useful. It certainly doesn't fit nicely with the template approach which is why I'm currently not using the declarations here.

2026/09/05: Same for /titleline (which I renamed to fragment because it may not be a "line"). Again I don't think it will ever differ for different levels.

```
\AssignStructureRole{sec/2/fragment}{heading-fragment}
\NewStructureName{sec/2/number}
\AssignStructureRole{sec/2/number}{heading-number}
\NewStructureName{sec/3}
\AssignStructureRole{sec/3}{subsubsection}
\NewStructureName{sec/3/title}
\AssignStructureRole{sec/3/title}{subsubsection-heading}
\NewStructureName{sec/3/fragment}
\AssignStructureRole{sec/3/fragment}{heading-fragment}
\NewStructureName{sec/3/number}
\AssignStructureRole{sec/3/number}{heading-number}
\NewStructureName{sec/4}
\AssignStructureRole{sec/4}{paragraph}
\NewStructureName{sec/4/title}
\AssignStructureRole{sec/4/title}{paragraph-heading}
\NewStructureName{sec/4/fragment}
\AssignStructureRole{sec/4/fragment}{heading-fragment}
\NewStructureName{sec/4/number}
\AssignStructureRole{sec/4/number}{heading-number}
\NewStructureName{sec/5}
\AssignStructureRole{sec/5}{subparagraph}
\NewStructureName{sec/5/title}
\AssignStructureRole{sec/5/title}{subparagraph-heading}
\NewStructureName{sec/5/fragment}
\AssignStructureRole{sec/5/fragment}{heading-fragment}
\NewStructureName{sec/5/number}
\AssignStructureRole{sec/5/number}{heading-number}
```

1.5.4 Table of contents and similar

```
\NewStructureName{toc}
\AssignStructureRole{toc}{TOC}
\NewStructureName{toc/item}
\AssignStructureRole{toc/item}{TOCI}
\NewStructureName{toc/item/label}
\AssignStructureRole{toc/item/label}{Lbl}
\NewStructureName{toc/item/text}
\AssignStructureRole{toc/item/text}{Reference}
\NewStructureName{toc/item/page}
\AssignStructureRole{toc/item/page}{Reference}
```

1.5.5 Title

"title" in this context can be a structure that surrounds the *title block* (or a part from it) with author, title, subtitle, date and perhaps more. Or it can mean only the *title text* (plus perhaps the subtitle). In PDF 2.0 the tag `Title` is described as grouping and block type and it can not be used in a `P` tag, so it is better used for the *title block*. With PDF 1.7 `Title` is currently rolemapped to `P` and the code in latex-lab-title uses it for the *title text*. We provide therefore two symbolic names for the two concepts, which both map to `Title`.

UFi:This should be discussed.

```
\NewStructureName{title/block}
\AssignStructureRole{title/block}{title-block}
\NewStructureName{title/text}
```

```
\AssignStructureRole{title/text}{title-text/latex}
\NewStructureName{title/author}
\AssignStructureRole{title/author}{title-author}
\NewStructureName{title/date}
\AssignStructureRole{title/date}{title-date}
```

1.5.6 Lists and blocks

`block/inner` is the generic named used for inner blocks in, e.g., the recipe `standard`.

```
\NewStructureName{block/inner}
\AssignStructureRole{block/inner}{Div}
\NewStructureName{block/itemize}
\AssignStructureRole{block/itemize}{itemize}
\NewStructureName{block/description}
\AssignStructureRole{block/description}{description}
\NewStructureName{block/enumerate}
\AssignStructureRole{block/enumerate}{enumerate}
\NewStructureName{block/list}
\AssignStructureRole{block/list}{list}
\NewStructureName{block/list/item}
\AssignStructureRole{block/list/item}{item}
\NewStructureName{block/list/label}
\AssignStructureRole{block/list/label}{itemlabel}
\NewStructureName{block/list/body}
\AssignStructureRole{block/list/body}{itembody}
```

```
\NewStructureName{block/quote}
\AssignStructureRole{block/quote}{quote}
```

decide on tagging for verse

For now we map `verse` to `quote` which means it will become BlockQuote in tagging.

```
\NewStructureName{block/verse}
\AssignStructureRole{block/verse}{quote}
```

```
\NewStructureName{block/quotation}
\AssignStructureRole{block/quotation}{quotation}
%
\NewStructureName{block/verbatim}
\AssignStructureRole{block/verbatim}{verbatim}
\NewStructureName{block/verbatim/codeline}
\AssignStructureRole{block/verbatim/codeline}{codeline}
\NewStructureName{block/verbatim/linenumber}
\AssignStructureRole{block/verbatim/linenumber}{Lbl}
%
\NewStructureName{block/theorem-like}
\AssignStructureRole{block/theorem-like}{theorem-like}
\NewStructureName{block/theorem-like/caption}
\AssignStructureRole{block/theorem-like/caption}{Caption}
\NewStructureName{block/theorem-like/label}
\AssignStructureRole{block/theorem-like/label}{Lbl}
```

For now we map proof to theorem-like:

```
\NewStructureName{block/proof}
\AssignStructureRole{block/proof}{theorem-like}
```

Bibliographies normally use internally lists and tag names should be adapted by changing the symbolic names of lists locally, so dedicated symbolic names are probably not needed.

```
%\NewStructureName{block/bib}
%\AssignStructureRole{block/bib}{L}
%\NewStructureName{block/bib/label}
%\AssignStructureRole{block/bib/label}{Lbl}
%\NewStructureName{block/bib/item}
%\AssignStructureRole{block/bib/item}{LI}
%\NewStructureName{block/bib/body}
%\AssignStructureRole{block/bib/body}{LBody}
```

1.5.7 LaTeX Code documentation

The class l3doc defines a number of structures for the document of LaTeX code. The exact tagging is currently experimental and it is unclear if symbolic names and tag names for this case should be setup here by default. Their setup is therefore done in latex-lab-l3doc-tagging.

1.5.8 References and Links

```
\NewStructureName{ref}
\AssignStructureRole{ref}{Reference}
\NewStructureName{link}
\AssignStructureRole{link}{Link}
\NewStructureName{cite}
\AssignStructureRole{cite}{Reference}
```

1.5.9 Tables

```
\NewStructureName{table}
\AssignStructureRole{table}{Table}
\NewStructureName{table/row}
\AssignStructureRole{table/row}{TR}
\NewStructureName{table/cell}
\AssignStructureRole{table/cell}{TD}
\NewStructureName{table/headercell}
\AssignStructureRole{table/headercell}{TH}
```

THead and TFoot are currently not used but are perhaps needed later

```
%\NewStructureName{tbl/table/head}
%\AssignStructureRole{tbl/table/head}{}
%\NewStructureName{tbl/table/foot}
%\AssignStructureRole{tbl/table/foot}{}
```

Inside p-cells we have to set the para/semantic structure to Div as Part is not allowed as child of TD and as we can not flatten the paragraphs.

```
\NewStructureName{table/para/semantic}
\AssignStructureRole{table/para/semantic}{Div}
```

We need also names for layout tables

```
\NewStructureName{layouttable}
\AssignStructureRole{layouttable}{Div}
```

```
\NewStructureName{layouttable/row}
\AssignStructureRole{layouttable/row}{NonStruct}
\NewStructureName{layouttable/cell}
\AssignStructureRole{layouttable/cell}{\UseStructureName{para/textblock}}
\NewStructureName{layouttable/pcell}
\AssignStructureRole{layouttable/pcell}{NonStruct}
```

1.5.10 Floats

```
\NewStructureName{float/generic}
\AssignStructureRole{float/generic}{float}
\NewStructureName{float/figures}
\AssignStructureRole{float/figures}{figures}
\NewStructureName{float/tables}
\AssignStructureRole{float/tables}{tables}
\NewStructureName{float/figure}
\AssignStructureRole{float/figure}{float}
\NewStructureName{float/table}
\AssignStructureRole{float/table}{float}
\NewStructureName{float/figure/caption}
\AssignStructureRole{float/figure/caption}{Caption}
\NewStructureName{float/table/caption}
\AssignStructureRole{float/table/caption}{Caption}
\NewStructureName{float/figure/label}
\AssignStructureRole{float/figure/label}{Lbl}
\NewStructureName{float/table/label}
\AssignStructureRole{float/table/label}{Lbl}
```

1.5.11 Footnotes

```
\NewStructureName{fnote/mark}
\AssignStructureRole{fnote/mark}{footnotemark}
\NewStructureName{fnote/note}
\AssignStructureRole{fnote/note}{footnote}
\NewStructureName{fnote/note/label}
\AssignStructureRole{fnote/note/label}{footnotelabel}
```

1.5.12 Graphics

```
\NewStructureName{graphic}
\AssignStructureRole{graphic}{Figure}
\NewStructureName{graphic/symbol}
\AssignStructureRole{graphic/symbol}{Span}
```

1.5.13 Marginnotes

```
\NewStructureName{marginnote}
\AssignStructureRole{marginnote}{Aside}
```

1.5.14 Minipages and parboxes

```
\NewStructureName{minipage}
\AssignStructureRole{minipage}{text-minipage}
\NewStructureName{parbox}
\AssignStructureRole{parbox}{text-parbox}
```

```
\NewStructureName{sec/parbox}
\AssignStructureRole{sec/parbox}{heading-parbox}
\NewStructureName{sec/minipage}
\AssignStructureRole{sec/minipage}{heading-minipage}
```

1.5.15 Text

TODO: These still need to be incorporated into the LaTeX namespace at some later time.

```
\NewStructureName{text/emph}
\AssignStructureRole{text/emph}{Em}
\NewStructureName{text/strong}
\AssignStructureRole{text/strong}{Strong}
\NewStructureName{text/verb}
\AssignStructureRole{text/verb}{Code}
```

1.5.16 Math

Math uses internally MathML and this should not be remapped so we need only names for the outer `Formula` tag and perhaps for a label:

```
\NewStructureName{math/inline}
\AssignStructureRole{math/inline}{Formula}
\NewStructureName{math/display}
\AssignStructureRole{math/display}{Formula}
\NewStructureName{math/label}
\AssignStructureRole{math/label}{Lbl}
```

1.5.17 Index and glossaries

Tagging not decided yet.

1.6 Attributes

The latex-lab code sets up a number of kernel attributes. These are collected here:

```
\tagpdfsetup
 {
```

1.7 Text attributes

These attributes are for super- and subscripts and are used in latex-lab-text:

```
   role/new-attribute = {sub} {/O /Layout /TextPosition/Sub},
   role/new-attribute = {sup} {/O /Layout /TextPosition/Sup},
```

These are for text alignment and they are used in latex-lab-block: role/new-attribute = justify /O /Layout /TextAlign/Justify, role/new-attribute = center /O /Layout /TextAlign/Center, role/new-attribute = raggedright/O /Layout /TextAlign/Start, role/new-attribute = raggedleft /O /Layout /TextAlign/End,

1.8 Math attributes

```
   role/new-attribute = {inline}  {/O /Layout /Placement/Inline},
   role/new-attribute = {display} {/O /Layout /Placement/Block},
```

1.9 List attributes

```
  role/new-attribute = {itemize}
                        {/O /List /ListNumbering/Unordered},
  role/new-attribute = {enumerate}
                        {/O /List /ListNumbering/Ordered},
  role/new-attribute = {description}
                        {/O /List /ListNumbering/Description},
```

Initially, we had `/None` for the basic `list` environment, but that is not allowed in PDF/UA-2 if the list contains any Lbl tags. So the `list` attribute uses `Unordered` too.

```
  role/new-attribute = {list}{/O /List /ListNumbering/Unordered},
```

1.10 Table attributes

The code predeclares three attribute names for the scope of header, `TH-col`, `TH-row`, `TH-both`. The latex-lab-table code declares on the fly additional attributes `colspan-`⟨*number*⟩ and `rowspan-`⟨*number*⟩ with the content `/O /Table /ColSpan` ⟨*number*⟩ and `/O /Table /RowSpan` ⟨*number*⟩

```
  role/new-attribute* = {TH-col}  {/O /Table /Scope /Column},
  role/new-attribute* = {TH-row}  {/O /Table /Scope /Row},
  role/new-attribute* = {TH-both} {/O /Table /Scope /Both},
  role/new-attribute = {ARIA-role-presentation}
                         {/O /ARIA-1.1/role (presentation)}

 } %end tagpdfsetup
```

⟨/package⟩