

The `latex-lab-table` package

Changes related to the tagging of tables

Frank & Ulrike, L^AT_EX Project*

v0.85t 2026-04-28

Abstract

The following code implements a first draft for the tagging of tables. It still has a large number of limitations and restrictions!

Contents

1	Documentation	2
1.1	Use	2
1.2	Untagged and presentation tables	2
1.3	Header rows and columns	3
1.4	Spanning of cells	3
2	Limitations	4
3	Introduction	5
4	Technical details and problems	5
4.1	TODOs	5
5	Implementation	5
5.1	Variables	5
5.2	Tagging support sockets	6
5.3	Environments	14
5.4	Interfaces to tagging	15
5.4.1	Tagging helper commands	15
5.4.2	Disabling/enabling	15
5.4.3	Header support	17
5.5	Multirow support	18
5.6	Misc stuff	20
5.7	longtable	22

*Initial implementation done by Frank Mittelbach

1 Documentation

In \LaTeX the word `table` is used as the name of the float environment that can contain a data table¹ along with a caption and some additional text. The environments for actual data tables have various names like `tabular`, `tabular*`, `tabularx` and `longtable`—the last should not be used inside a float and supports its own caption command.

In this documentation “table” always means such data tables and not the float environment.

From a high level perspective the tagging of tables doesn’t look very difficult: one only has to surround rows and cells by `TR` and `TH` or `TD` structures. But there are difficulties:

- One is that over the years various packages related to tables have been written that all change some of the internals. Inserting the tagging commands and testing all the variants and various nestings is not trivial.
- The other difficulty is that most of the existing environments to create tables do not know the concept of headers as a semantic structure. Headers are only produced through visual formatting, e.g., by making them bold or by underlying some color. But accessible tables should have properly marked up headers and this means that additional syntax to declare headers (when they can’t be guessed) must be developed. This is still an area for research.

Right now, this module therefore implements only some basic support for the tagging of tables. A list of the known limitations is shown below.

1.1 Use

The module is loaded when you use `\DocumentMetadata` and if tagging is active it will automatically tag all already supported table environments the exception of tables in the header and footer of the page (where tagging is disabled). Such tables can be nested.

Inside cells the automatic tagging of paragraphs is disabled with the exception of `p/m/b`-type cells.

Rows do not need to contain a full number of `&`, missing cells are automatically added with an empty `TD`-structure.

You should not insert meaningful text with `!\{...\}` or `@{...\}` or `\noalign` between the rows or columns of the table. With `pdflatex` such text will be unmarked, with `lualatex` it will be marked as artifact. Either case means that it will be currently ignored in the structure.²

As mentioned below the `colortbl` doesn’t work fully with the tagging, but when it does, colors inside the table are currently ignored by the tagging. If such a color has a semantic meaning (like “important value”) this meaning will be lost.

1.2 Untagged and presentation tables

If a table should not be tagged as table, for example because it is merely used as a layout tool to ensure that the content is properly aligned or because it is a not yet (fully) supported table structure, the tagging can be disabled with `\tagpdfsetup{table/tagging=false}`

¹But it does not really have to, you can put other material into such environments.

²While it is theoretically possible to collect such text and move it into a structure it would require manual markup from the author to clarify where this text belongs too.

or changed with `\tagpdfsetup{table/tagging=presentation}` or `\tagpdfsetup{table/tagging=div}`.

³ The first option disables the table tagging code and the content of the tabular is then treated more or less like running text. This works ok for simple tables using only hmode-cells (l/c/r) with normal text content, but fails if the table uses vmode-cells (p/m/b). In such cases the second option works better: it keeps the tagging code active, but adds an ARIA-role as attribute to indicate that this is a presentation table. When deriving to html this gives the best result as it keeps the layout. The last option changes the tag names and creates simply a number of DIV structures. Both options also (re)set the `table/header-rows` and `table/header-columns` keys to empty.

To reset to the normal tagging in the current group (e.g. inside a presentation table) one can use `\tagpdfsetup{table/tagging=true}` (and if needed reenale the headers).

1.3 Header rows and columns

There is some basic support⁴ for headers. With

```
\tagpdfsetup{table/header-rows={\list of row numbers}}
\tagpdfsetup{table/header-columns={\list of column numbers}}
```

you can declare which (absolute) row and column numbers should be tagged as header rows and header columns. It applies to all tables until it is changed to a different list of row or column numbers or undone by setting the keys to `\empty`. A row or column number can be negative, then the counting starts from the end of the table. In a `longtable` the code will currently use the `\endhead` or `\endfirsthead` rows as header if one of these commands has been used and in that case the code ignores a `table/header-rows` setting.

1.4 Spanning of cells

`\multicolumn` is supported out of the box and will create a structure with a suitable `ColSpan` attribute⁵

For cells spanning rows some preliminary support exists⁶: If you add

```
\tagpdfsetup{table/multirow={\number of rows}}
```

to a cell the code will add the suitable `RowSpan` attribute and suppress the tagging of affected cells in the following rows. This will also work if the current cell is a `\multicolumn`, then the cell will span both rows and columns. It is the duty of the author to ensure that all cells below and covered by a `RowSpan` are empty! The code neither checks that nor does it tries to suppress content.

Feedback and problems with the code can be reported at <https://github.com/latex3/tagging-project> either in form of explicit issues or as a “discussion topic”, whatever works best.

³The value `presentation` refers to the ARIA role “presentation”.

⁴This is not meant to be the final interface, though.

⁵The code uses actually an attribute `class`. The validator PAC doesn’t handle this correctly currently and complains about a missing attribute.

⁶The interface is bound to change!

2 Limitations

- The code loads the `array` package and so does not work without it (that is not really a limitation, but can affect existing tables).
- It supports only a restricted number of tables types. Currently `tabular`, `tabular*`, `tabularx`, and `longtable`.
- The `array` environment is assumed to be part of math and tagging as a table is disabled for it.
- Row spans can only be added manually (the `multirow` package is untested).
- The `colortbl` package breaks tagging if there are nested tables.
- The `tabularray` package uses a completely different method to create tables and is not supported by this code.
- Tables in text mode made with `nicematrix` package are often correctly tagged, but should be carefully checked. In math mode they are currently not compatible with the MathML structure elements.
- Most other packages related to tables in \LaTeX are not fully tested; that includes packages that change rules like `booktabs`, `hhline`, `arydshln`, `hvdashln`. Some problems have been resolved, either in the packages or through a firstaid.
- `longtable` currently only works with `lualatex`. With other engines it breaks as its output routine clashes with the code which closes open MC-chunks at pagebreaks and if this is resolved there will probably be problems with the head and foot boxes (but this can't be tested currently).
- Not every table should be tagged as a Table structure, often they are only used as layout help, e.g. to align authors in title pages. In such uses the tagging of the table must be deactivated with `\tagpdfsetup{table/tagging=false}` or changed with `\tagpdfsetup{table/tagging=presentation}` or `\tagpdfsetup{table/tagging=div}`.
- Only simple header rows and columns are currently supported. Complex headers with subheaders will be handled later as that needs some syntax changes. Tables with more than one header row or column are probably not pdf/UA conformant as the headers array in the cells is missing.
- A `longtable` `\caption` is currently simply formatted as a multicolumn and not tagged as a `Caption` structure.
- The `caption` package will quite probably break the `longtable` caption.
- The setup for `longtable` requires lots of patches to internal `longtable` commands and so can easily break if other packages try to patch `longtable` too.
- The `longtable` environment supports footnotes in p-type columns, but it hasn't been tested yet if this works also with the tagging code.
- Vertical boxes (`\parbox`, `minipage`, ...) inside cells can be problematic.
- The code is quite noisy and fills the log with lots of messages.⁷

⁷Helpful for us at this stage.

3 Introduction

4 Technical details and problems

The implementation has to take care of various details.

4.1 TODOs

- Test `array-006-longtable.lvt` and `array-007-longtable.lvt` have errors with pdfTeX (para tagging)
- Instead of before/after hooks we should add sockets directly into the code.
- Debugging code and messages must be improved.
- Cells need an `Headers` array.
- Row spans should be supported (but perhaps need syntax support)
- Longtable captions should be properly supported.
- More packages must be tested.

5 Implementation

```
1 <@@=tbl>
2 <*package>

3 \ProvidesExplPackage {latex-lab-testphase-table} {\ltabletbldate} {\ltabletblversion}
4 {Code related to the tagging of tables}
```

This builds on `array` so we load it by default:

```
5 \RequirePackage{array}
```

5.1 Variables

This is for the celltag, e.g. TD or TH:

```
6 \tl_new:N \l__tbl_celltag_tl
7 \tl_set:Nn \l__tbl_celltag_tl {\UseStructureName{table/cell}}
8 \tl_new:N \l__tbl_pcelltag_tl
9 \tl_set:Nn \l__tbl_pcelltag_tl {\UseStructureName{table/cell}}
```

For the rowtag, probably always TR:

```
10 \tl_new:N \l__tbl_rowtag_tl
11 \tl_set:Nn \l__tbl_rowtag_tl {\UseStructureName{table/row}}
```

For the tabletag, probably always Table:

```
12 \tl_new:N \l__tbl_tabletag_tl
13 \tl_set:Nn \l__tbl_tabletag_tl {\UseStructureName{table}}
```

And here cell and row attributes:

```

14 \tl_new:N \l__tbl_cellattribute_tl
15 \tl_set:Nn \l__tbl_cellattribute_tl {}
16 \tl_new:N \l__tbl_rowattribute_tl
17 \tl_set:Nn \l__tbl_rowattribute_tl {}

```

Variable to store cell info like which cell must be skipped when tagging multirows.

```

18 \prop_new:N\g__tbl_untagged_cells_prop
19 \prop_new:N\l__tbl_saved_untagged_cells_prop

```

Temp variables

```

20 \clist_new:N \l__tbl_tmpa_clist
21 \tl_new:N \l__tbl_tmpa_tl
22 \str_new:N \l__tbl_tmpa_str

```

(End of definition for \l__tbl_celltag_tl \l__tbl_pcelltag_tl and others.)

5.2 Tagging support sockets

These are the standard plugs for tagging of cells and rows.

The `tbl/init/celldata` and `tbl/restore/celldata` are defined in `ltagging`.

default (plug) This socket is used inside `\tbl_init_cell_data_for_table` for the case that this a nested table in a cell.

```

23 \NewTaggingSocketPlug{tbl/init/celldata}{default}
24 {
25   \prop_set_eq:NN\l__tbl_saved_untagged_cells_prop \g__tbl_untagged_cells_prop
26   \prop_gclear:N \g__tbl_untagged_cells_prop
27 }
28 \AssignTaggingSocketPlug{tbl/init/celldata}{default}

```

default (plug) This socket is used inside `\tbl_restore_outer_cell_data:` and restores values when a nested table is ended.

```

29 \NewTaggingSocketPlug{tbl/restore/celldata}{default}
30 {
31   \prop_gset_eq:NN\g__tbl_untagged_cells_prop \l__tbl_saved_untagged_cells_prop
32 }
33 \AssignTaggingSocketPlug{tbl/restore/celldata}{default}

```

TD (plug)

```

34 \NewTaggingSocketPlug{tbl/cell/begin}{TD}
35 {

```

Next line was previously outside of the plug, so if we want to execute it always even if the noop plug is in force this needs a different solution.

```

36   \__tbl_show_curr_cell_data:

```

We test if the cell should be tagged at all.

```

37   \__tbl_if_tag_cell:nnT {\g__tbl_row_int } { \g__tbl_col_int }
38   {

```

this sets row headers

```

39         \clist_if_in:NVT \l__tbl_header_columns_clist\g__tbl_col_int
40         {
41             \tl_set:Ne \l__tbl_celltag_tl {\UseStructureName{table/headercell}}
42             \tl_set:Ne \l__tbl_cellattribute_tl {\l__tbl_cellattribute_tl,TH-row}
43         }

```

Here we handle negative value for row headers:

```

44         \tl_set:Ne \l__tbl_tmpa_tl
45         {\int_eval:n { \g__tbl_col_int - 1 - \g__tbl_table_cols_tl }}
46         \clist_if_in:NoT \l__tbl_header_columns_clist { \l__tbl_tmpa_tl }
47         {
48             \tl_set:Ne \l__tbl_celltag_tl {\UseStructureName{table/headercell}}
49             \tl_set:Ne \l__tbl_cellattribute_tl {\l__tbl_cellattribute_tl,TH-row}
50         }
51         \tag_struct_begin:n
52         {
53             tag                =\l__tbl_celltag_tl,
54             attribute-class    ={\l__tbl_cellattribute_tl}
55         }
56         \seq_gput_right:Ne \g__tbl_struct_cur_seq { \tag_get:n {struct_num} }

```

we store the cells of multicolumns as negative number. This allow to skip them or to use them as needed.

```

57         \int_step_inline:nn { \g__tbl_span_tl - 1 }
58         {
59             \seq_gput_right:Ne \g__tbl_struct_cur_seq { -\tag_get:n {struct_num} }
60         }
61         \tag_mc_begin:n{ }
62     }
63 }

```

TD (*plug*)

```

64 \NewTaggingSocketPlug{tbl/cell/end}{TD}
65 {
66     \__tbl_if_tag_cell:nnT {\g__tbl_row_int } { \g__tbl_col_int }
67     {
68         \tag_mc_end:
69         \tag_struct_end:
70     }
71 }

```

In p-columns we need a slightly different plug which reactivates the paragraph tagging.

TDpbox (*plug*)

```

72 \NewTaggingSocketPlug{tbl/pcell/begin}{TDpbox}
73 {
74     \__tbl_show_curr_cell_data:
75     \__tbl_if_tag_cell:nnT {\g__tbl_row_int } { \g__tbl_col_int }
76     {
77         \clist_if_in:NVT \l__tbl_header_columns_clist\g__tbl_col_int
78         {
79             \tl_set:Ne \l__tbl_pcelltag_tl {\UseStructureName{table/headercell}}

```

```

80         \tl_set:Nc \l__tbl_cellattribute_tl {\l__tbl_cellattribute_tl,TH-row}
81     }
82     \tag_struct_begin:n
83     {
84         tag                =\l__tbl_pcelltag_tl,
85         attribute-class ={\l__tbl_cellattribute_tl}
86     }
87     \seq_gput_right:Nc \g__tbl_struct_cur_seq { \tag_get:n {struct_num} }
88     \int_step_inline:nn { \g__tbl_span_tl - 1 }
89     {
90         \seq_gput_right:Nc \g__tbl_struct_cur_seq { -\tag_get:n {struct_num} }
91     }
92     \tagpdfparaOn
93     \tl_set:Nn \l__tbl_tag_para_main_tag_tl {\UseStructureName{table/para/semantic}}
94 }
95 }

```

TDpbox (*plug*)

```

96 \NewTaggingSocketPlug{tbl/pcell/end}{TDpbox}
97 {
98     \__tbl_if_tag_cell:nnT {\g__tbl_row_int } { \g__tbl_col_int }
99     {
100         \tag_struct_end:
101         \legacy_if:nT {@endpe}{\par}
102         \mode_if_vertical:T{ \tagpdfparaOff }
103     }
104 }

```

TR (*plug*)

```

105 \NewTaggingSocketPlug{tbl/row/begin}{TR}
106 {
107     \seq_gclear:N \g__tbl_struct_cur_seq
108     \tag_struct_begin:n
109     {
110         tag                =\l__tbl_rowtag_tl,
111         attribute-class=\l__tbl_rowattribute_tl
112     }
113     \seq_gput_right:Nc \g__tbl_struct_rows_seq { \tag_get:n {struct_num} }
114 }

```

TR (*plug*)

```

115 \NewTaggingSocketPlug{tbl/row/end}{TR}
116 {
117     \__tbl_add_missing_cells:
118     \seq_gput_right:Nc \g__tbl_struct_cells_seq
119     {
120         \seq_use:Nn \g__tbl_struct_cur_seq {,}
121     }
122     \int_compare:nNnTF { \g__tbl_row_int } =
123     { \seq_count:N \g__tbl_struct_cells_seq }
124     {
125         \__tbl_trace:n
126         {==>~
127             structure~stored~for~row~\int_use:N\g__tbl_row_int :~

```



```

128         \seq_use:Nn \g__tbl_struct_cur_seq {,}
129     }
130 }
131 { \ERRORtbl/row } % should not happen ...
132 \tag_struct_end:
133 }

```

And the plugs for the table as whole. The code can be different for normal tables which can also be used inline and nested and “vmode” tables like longtable.

Table (plug) Inside a TD or TR Part is not allowed as child. For structures that restore paragraph settings we therefore need a special plug that adjust the settings. Currently we set the para main tag to Div. This leads to double Div-structures but flattening the paragraph doesn’t work, it errors if there is a list inside.

```

134 \NewTaggingSocketPlug{para/restore}{Table}
135 {
136     \tl_set:Nn \l__tag_para_main_tag_tl {\UseStructureName{table/para/semantic}}
137     \tl_set_eq:NN \l__tag_para_tag_tl \l__tag_para_tag_default_tl
138     \bool_set_true:N \l__tag_para_bool
139     \bool_set_false:N \l__tag_para_flattened_bool
140     \int_zero:N \l__tag_block_flattened_level_int
141 }

```

Table (plug) Inside a table we currently only disable paratagging. We assume that these sockets are called in a group, so there is no need to reenale paratagging.

```

142 \NewTaggingSocketPlug{tbl/init}{Table}
143 {
144     \bool_set_false:N \l__tag_para_bool
145     \AssignTaggingSocketPlug{para/restore}{Table}

```

We also initialize the structure data variables a this point.

```

146     \__tbl_init_struct_data:
147 }

```

Table (plug) This plug will fine tune the structure.

```

148 \NewTaggingSocketPlug{tbl/finalize}{Table}
149 {
150     \__tbl_set_header_rows:

```

Similarly, we restore the outer values of the structure data when we leave the table.

```

151     \__tbl_restore_struct_data:
152 }

```

Table (plug) This plug will initialize the structure in longtable.

```

153 \NewTaggingSocketPlug{tbl/longtable/init}{Table}
154 {
155     \seq_gclear:N\g__tbl_struct_rows_seq
156     \seq_gclear:N\g__tbl_struct_cells_seq
157     \seq_gclear:N\g__tbl_struct_cur_seq
158     \seq_gclear:N\g__tbl_LT@firsthead_rows_seq
159     \seq_gclear:N\g__tbl_LT@head_rows_seq
160     \seq_gclear:N\g__tbl_LT@lastfoot_rows_seq
161     \seq_gclear:N\g__tbl_LT@foot_rows_seq
162 }

```

Table (*plug*) This plug will fine tune the structure in longtable.

```
163 \NewTaggingSocketPlug{tbl/longtable/finalize}{Table}
164 {
```

If neither `\endhead` nor `\endfirsthead` has been used we use the standard header command:

```
165     \bool_lazy_and:nnTF
166     { \seq_if_empty_p:N \g__tbl_LT@head_rows_seq }
167     { \seq_if_empty_p:N \g__tbl_LT@firsthead_rows_seq }
168     { __tbl_set_header_rows: }
```

Otherwise, if `firsthead` has not been used we use `head`. For this we simple retrieve the row numbers and then call the header command.

```
169     {
170     \seq_if_empty:NNTF \g__tbl_LT@firsthead_rows_seq
171     {
172         \clist_set:Nc \l__tbl_header_rows_clist
173         { \seq_use:Nn \g__tbl_LT@head_rows_seq {,} }
174         __tbl_set_header_rows:
175     }
```

In the other case we use `firsthead`.

```
176     {
177         \clist_set:Nc \l__tbl_header_rows_clist
178         { \seq_use:Nn \g__tbl_LT@firsthead_rows_seq {,} }
179         __tbl_set_header_rows:
```

Additionally we have to remove the head to avoid duplication. The one option here is to remove the rows from the kid sequence of the table (which will lead to orphaned structure elements), the other to make them artifact. For now we use the first option for pdf 1.7 and the second for pdf 2.0.

```
180         \pdf_version_compare:NnTF < {2.0}
181         {
182             \seq_map_inline:Nn \g__tbl_LT@head_rows_seq
183             {
184                 \seq_gset_item:cnn
185                 {g__tag_struct_kids_ \g__tbl_struct_table_t1 _seq}
186                 { ##1 }
187             }
```

Not sure if needed, but if needed we can remove also the P tag. This is currently disabled as it produce warnings. TODO: This needs also a `tagpdf` command which takes care of debug code.

```
188         \tl_set:Nc \l__tbl_tmpa_tl
189         { \seq_item:Nn \g__tbl_struct_rows_seq {##1} }
190         \prop_if_exist:cT
191         { g__tag_struct_ \l__tbl_tmpa_tl _prop }
192         {
193             %\prop_gremove:cn {g__tag_struct_ \l__tbl_tmpa_tl _prop} {P}
194         }
195     }
196 }
197 {
```

```

198         \seq_map_inline:Nn \g__tbl_LT@head_rows_seq
199         {
200             \tl_set:Ne \l__tbl_tmpa_tl
201             { \seq_item:Nn \g__tbl_struct_rows_seq {##1} }
202             \prop_if_exist:cT
203             { g__tag_struct_ \l__tbl_tmpa_tl _prop }
204             {
205                 \__tag_struct_prop_gput:onn { \l__tbl_tmpa_tl } {S}{/Artifact}
206             }
207         }
208     }
209 }
210

```

The foot is handled similar, the difference is that we have to move it to the end one of them is not empty, but do nothing if they aren't there.

```

211 \bool_lazy_and:nnF
212 { \seq_if_empty_p:N \g__tbl_LT@foot_rows_seq }
213 { \seq_if_empty_p:N \g__tbl_LT@lastfoot_rows_seq }
214 {

```

If lastfoot is empty move foot to the end.

```

215     \seq_if_empty:NTF \g__tbl_LT@lastfoot_rows_seq
216     {
217         \seq_map_inline:Nn \g__tbl_LT@foot_rows_seq
218         {
219             \tl_set:Ne \l__tbl_tmpa_tl
220             {
221                 \seq_item:cn
222                 {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}
223                 {##1}
224             }
225             \seq_gset_item:cnn
226             {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}
227             { ##1 }
228             {}
229             \seq_gput_right:cV
230             {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}
231             \l__tbl_tmpa_tl
232         }
233     }

```

If lastfoot is not empty we move that.

```

234     {
235         \seq_map_inline:Nn \g__tbl_LT@lastfoot_rows_seq
236         {
237             \tl_set:Ne \l__tbl_tmpa_tl
238             {
239                 \seq_item:cn
240                 {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}
241                 {##1}
242             }
243             \seq_gset_item:cnn
244             {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}

```

```

245         { ##1 }
246     {}
247     \seq_gput_right:cV
248     {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}
249     \l__tbl_tmpa_tl
250 }

```

and we hide foot

```

251     \pdf_version_compare:NnTF < {2.0}
252     {
253         \seq_map_inline:Nn \g__tbl_LT@foot_rows_seq
254         {
255             \seq_gset_item:cnn
256             {g__tag_struct_kids_ \g__tbl_struct_table_tl _seq}
257             { ##1 }
258         }

```

Not sure if needed, but if needed we can remove also the P tag. This is currently disabled as it produce warnings. TODO: This needs also a tagpdf command which takes care of debug code.

```

259         \tl_set:Ne \l__tbl_tmpa_tl
260         { \seq_item:Nn\g__tbl_struct_rows_seq {##1} }
261         \prop_if_exist:cT
262         { g__tag_struct_ \l__tbl_tmpa_tl _prop }
263         {
264             %\prop_gremove:cn {g__tag_struct_ \l__tbl_tmpa_tl _prop} {P}
265         }
266     }
267 }
268 {
269     \seq_map_inline:Nn \g__tbl_LT@foot_rows_seq
270     {
271         \tl_set:Ne \l__tbl_tmpa_tl
272         { \seq_item:Nn\g__tbl_struct_rows_seq {##1} }
273         \prop_if_exist:cT
274         { g__tag_struct_ \l__tbl_tmpa_tl _prop }
275         {
276             \__tag_struct_prop_gput:onn {\l__tbl_tmpa_tl} {S}{/Artifact}
277         }
278     }
279 }
280 }
281 }
282 }

```

Table (plug)

We must avoid that the reuse of the header foot box leads to duplicated content, thus reset attribute of the box:

```

283 \NewTaggingSocketPlug{tbl/longtable/head}{Table}
284 {
285     \tagmcbegin{artifact}
286     \tag_mc_reset_box:N\LT@head
287     \tagmcend
288 }

```

Table (*plug*)

```

289 \NewTaggingSocketPlug{tbl/longtable/foot}{Table}
290 {
291   \tagmcbegin{artifact}
292   \tag_mc_reset_box:N \LT@foot
293   \tagmccend
294 }

```

Table (*plug*)

```

295 \NewTaggingSocketPlug{tbl/hmode/begin}{Table}
296 {
297   \tag_mc_end_push:

```

Close the P-chunk. This assumes that para-tagging is active. For nested tables that is not necessarily true, so we test for it.

```

298   \bool_lazy_and:nnT
299   { \bool_if_exist_p:N \l__tag_para_bool } { \l__tag_para_bool }
300   { \tag_struct_end:n { text } }
301   \tag_struct_begin:n {tag=\l__tbl_tabletag_tl}
302   \tl_gset:Ne \g__tbl_struct_table_tl { \tag_get:n {struct_num} }
303 }

```

LayoutTable (*plug*) This plug is meant for presentation tables, it sets the ARIA role.

```

304 \NewTaggingSocketPlug{tbl/hmode/begin}{LayoutTable}
305 {
306   \tag_mc_end_push:

```

Close the P-chunk. This assumes that para-tagging is active. For nested tables that is not necessarily true, so we test for it.

```

307   \bool_lazy_and:nnT
308   { \bool_if_exist_p:N \l__tag_para_bool } { \l__tag_para_bool }
309   { \tag_struct_end:n { text } }
310   \tag_struct_begin:n {tag=\l__tbl_tabletag_tl,attribute-class=ARIA-role-presentation}
311   \tl_gset:Ne \g__tbl_struct_table_tl { \tag_get:n {struct_num} }
312 }

```

Table (*plug*)

```

313 \NewTaggingSocketPlug{tbl/hmode/end}{Table}
314 {
315   \tag_struct_end:

```

reopen the P-chunk. This assumes that para-tagging is active. For nested tables that is not necessarily true, so we test for it.

```

316   \bool_lazy_and:nnT
317   { \bool_if_exist_p:N \l__tag_para_bool } { \l__tag_para_bool }
318   { \tag_struct_begin:n { tag=\UseStructureName{para/textblock} } }
319   \tag_mc_begin_pop:n{}
320 }

```

Table (*plug*)

```

321 \NewTaggingSocketPlug{tbl/vmode/begin}{Table}
322 {
323   \tag_struct_begin:n {tag=\l__tbl_tabletag_tl}
324   \tl_gset:Ne \g__tbl_struct_table_tl { \tag_get:n {struct_num} }
325 }

```

Table (*plug*)

```

326 \NewTaggingSocketPlug{tbl/vmode/begin}{LayoutTable}
327 {
328   \tag_struct_begin:n {tag=\l__tbl_tabletag_tl,attribute-class=ARIA-role=presentation}
329   \tl_gset:Ne \g__tbl_struct_table_tl { \tag_get:n {struct_num} }
330 }

```

Table (*plug*)

```

331 \NewTaggingSocketPlug{tbl/vmode/end}{Table}
332 {
333   \tag_struct_end:
334   \par
335 }

```

code (*plug*) This socket takes a number, checks if is larger than one, checks if the colspan attribute already exists (we can't predefine an arbitrary number), and updates \l__tbl_cellattribute_tl.

```

336 \NewTaggingSocketPlug{tbl/colspan}{code}
337 {
338   \int_compare:nNnT {#1}>{1}
339   {
340     \prop_get:NeNF \g__tag_attr_entries_prop
341       {colspan-\int_eval:n{#1}}
342     \l__tbl_tmpa_tl
343     {
344       \__tag_attr_new_entry:ee
345       {colspan-\int_eval:n{#1}}
346       {/0 /Table /ColSpan-\int_eval:n{#1}}
347     }
348     \tl_set:Ne \l__tbl_cellattribute_tl
349       {\l__tbl_cellattribute_tl,colspan-\int_eval:n{#1}}
350   }
351 }

```

code (*plug*) The sockets are declared in ltagging.dtx.

```

352 \NewTaggingSocketPlug{tbl/leaders/begin}{code}
353 { \tag_mc_begin:n{artifact} }
354 \NewTaggingSocketPlug{tbl/leaders/end}{code}
355 { \tag_mc_end: }

```

5.3 Environments

Currently we support only `tabular`, `tabular*`, `tabularx` and `longtable` (and possibly environments build directly on top of them).

The `array` environment is math. So we disable table tagging for now. We use the command hook to catch also cases where `\array` is used directly in other environments like matrix environments. Perhaps table tagging should be disable for math generally, but then we have to handle text insertions.

```

356 \AddToHook{cmd/array/before}{\__tag_tbl_disable:}

```

5.4 Interfaces to tagging

5.4.1 Tagging helper commands

5.4.2 Disabling/enabling

For now we have only the option true/false but this will probably be extended to allow different setups like first row header etc.

`__tag_tbl_disable:`

```
357 \cs_new_protected:Npn \__tag_tbl_disable:
358 {
359   \AssignTaggingSocketPlug{tbl/cell/begin}{noop}
360   \AssignTaggingSocketPlug{tbl/cell/end}{noop}
361   \AssignTaggingSocketPlug{tbl/pcell/begin}{noop}
362   \AssignTaggingSocketPlug{tbl/pcell/end}{noop}
363   \AssignTaggingSocketPlug{tbl/row/begin}{noop}
364   \AssignTaggingSocketPlug{tbl/row/end}{noop}
365   \AssignTaggingSocketPlug{tbl/init}{noop}
366   \AssignTaggingSocketPlug{tbl/finalize}{noop}
367   \AssignTaggingSocketPlug{tbl/longtable/init}{noop}
368   \AssignTaggingSocketPlug{tbl/longtable/head}{noop}
369   \AssignTaggingSocketPlug{tbl/longtable/foot}{noop}
370   \AssignTaggingSocketPlug{tbl/longtable/finalize}{noop}
371   \AssignTaggingSocketPlug{tbl/hmode/begin}{noop}
372   \AssignTaggingSocketPlug{tbl/hmode/end}{noop}
373   \AssignTaggingSocketPlug{tbl/vmode/begin}{noop}
374   \AssignTaggingSocketPlug{tbl/vmode/end}{noop}
375   \AssignTaggingSocketPlug{tbl/colspan}{noop}
376   \AssignTaggingSocketPlug{tbl/leaders/begin}{noop}
377   \AssignTaggingSocketPlug{tbl/leaders/end}{noop}
378 }
```

(End of definition for __tag_tbl_disable:.)

`__tag_tbl_enable:`

```
379 \cs_new_protected:Npn \__tag_tbl_enable:
380 {
381   \AssignTaggingSocketPlug{tbl/cell/begin}{TD}
382   \AssignTaggingSocketPlug{tbl/cell/end}{TD}
383   \AssignTaggingSocketPlug{tbl/pcell/begin}{TDpbox}
384   \AssignTaggingSocketPlug{tbl/pcell/end}{TDpbox}
385   \AssignTaggingSocketPlug{tbl/row/begin}{TR}
386   \AssignTaggingSocketPlug{tbl/row/end}{TR}
387   \AssignTaggingSocketPlug{tbl/init}{Table}
388   \AssignTaggingSocketPlug{tbl/finalize}{Table}
389   \AssignTaggingSocketPlug{tbl/longtable/head}{Table}
390   \AssignTaggingSocketPlug{tbl/longtable/foot}{Table}
391   \AssignTaggingSocketPlug{tbl/longtable/init}{Table}
392   \AssignTaggingSocketPlug{tbl/longtable/finalize}{Table}
393   \AssignTaggingSocketPlug{tbl/hmode/begin}{Table}
394   \AssignTaggingSocketPlug{tbl/hmode/end}{Table}
395   \AssignTaggingSocketPlug{tbl/vmode/begin}{Table}
396   \AssignTaggingSocketPlug{tbl/vmode/end}{Table}
397   \AssignTaggingSocketPlug{tbl/colspan}{code}
```

```

398 \AssignTaggingSocketPlug{tbl/leaders/begin}{code}
399 \AssignTaggingSocketPlug{tbl/leaders/end}{code}
400 }

```

(End of definition for `__tag_tbl_enable:`.)

```

\__tbl_init_tagnames_default: These commands are used to switch the tagnames.
\__tbl_init_tagnames_div:
401 \cs_new_protected:Npn\__tbl_init_tagnames_default:
402 {
403   \tl_set:Nn\l__tbl_rowtag_tl {\UseStructureName{table/row}}
404   \tl_set:Nn\l__tbl_pcelltag_tl {\UseStructureName{table/cell}}
405   \tl_set:Nn\l__tbl_celltag_tl {\UseStructureName{table/cell}}
406   \tl_set:Nn\l__tbl_tabletag_tl {\UseStructureName{table}}
407 }
408
409 \cs_new_protected:Npn\__tbl_init_tagnames_div:
410 {
411   \tl_set:Nn\l__tbl_rowtag_tl {\UseStructureName{layouttable/row}}
412   \tl_set:Nn\l__tbl_pcelltag_tl {\UseStructureName{layouttable/pcell}}
413   \tl_set:Nn\l__tbl_celltag_tl {\UseStructureName{layouttable/cell}}
414   \tl_set:Nn\l__tbl_tabletag_tl {\UseStructureName{layouttable}}
415 }

```

(End of definition for `__tbl_init_tagnames_default:` and `__tbl_init_tagnames_div:`.)

```

416 % See tagpdfsetup-keys.md in tagpdf/doc for the naming scheme.
417 \keys_define:nn { __tag / setup }
418 {
419   table/tagging .choices:nn = { true, on }
420   {
421     \__tag_tbl_enable:
422     \__tbl_init_tagnames_default:
423   },
424   table/tagging .choices:nn = { false, off }
425   { \__tag_tbl_disable: },
426   table/tagging .choice:,
427   table/tagging / presentation .code:n =
428   {
429     \__tag_tbl_enable:
430     \__tbl_init_tagnames_default:
431     \AssignTaggingSocketPlug{tbl/hmode/begin}{LayoutTable}
432     \AssignTaggingSocketPlug{tbl/vmode/begin}{LayoutTable}
433     \clist_clear:N \l__tbl_header_rows_clist
434     \clist_clear:N \l__tbl_header_columns_clist
435   },
436   table/tagging / div .code:n =
437   {
438     \__tag_tbl_enable:
439     \__tbl_init_tagnames_div:
440     \AssignTaggingSocketPlug{tbl/hmode/begin}{LayoutTable}
441     \AssignTaggingSocketPlug{tbl/vmode/begin}{LayoutTable}
442     \clist_clear:N \l__tbl_header_rows_clist
443     \clist_clear:N \l__tbl_header_columns_clist
444   },
445   table/tagging .default:n = true,

```



```

446     table/tagging .initial:n = true
447 }

```

This are the old key names kept for now for compatibility. They will got at some time.

```

448 \keys_define:nn { __tag / setup }
449 {
450     table-tagging .meta:n = {table/tagging={#1}}
451 }

```

5.4.3 Header support

Accessible table must have header cells declaring the meaning of the data in a row or column. To allow a data cell to find it header cell(s) a number of things must be done:

- every cell meant as a header should use the tag `TH`.
- header cells should have a `Scope` attribute with the value `Column`, `Row` or `Both`. This is not needed in the first row or column of a table.
- For more complex cases both `TD` and `TH` cell can contain a `Headers` attribute, that is an array of IDs of `TH` cell.

For now we support only header rows.

The attributes `TH-col`, `TH-row` and `TH-both` (for the scope of a header column) and `ARIA-role-presentation` are declared in `latex-lab-namespaces`.

```

\l__tbl_header_rows_clist
\l__tbl_header_columns_clist

```

This holds the numbers of the header rows and columns. Negative numbers are possible and count from the last column backwards.

```

452 \clist_new:N \l__tbl_header_rows_clist
453 \clist_new:N \l__tbl_header_columns_clist

```

```

\__tbl_set_header_rows:

```

```

454 \cs_new_protected:Npn \__tbl_set_header_rows:
455 {
456     \clist_map_inline:Nn \l__tbl_header_rows_clist
457     {
458         \clist_set:N\l__tbl_tmpa_clist
459         {
460             \seq_item:Nn \g__tbl_struct_cells_seq {##1}
461         }
462         \clist_map_inline:Nn \l__tbl_tmpa_clist
463         {

```

We can have negative numbers in the list from the multicolumn.

```

464         \prop_if_exist:cT { g__tag_struct_####1_prop }
465         {
466             \__tag_struct_prop_gput:nne{ ####1 }{S}{\UseStructureName{table/headercell}}

```

This sets the scope class. If the header has already a row attribute we replace by TH-both.

```

467 \prop_get:cnNTF
468 { g__tag_struct_####1_prop }
469 { C }
470 \l__tbl_tmpa_tl
471 {
472   \str_set:Ne \l__tbl_tmpa_str {\l__tbl_tmpa_tl}
473   \str_remove_once:Nn \l__tbl_tmpa_str {}
474   \str_remove_once:Nn \l__tbl_tmpa_str {}
475   \str_if_in:NnTF\l__tbl_tmpa_str{/TH-row}
476   {
477     \str_replace_once:Nnn \l__tbl_tmpa_str {/TH-row}{/TH-both}
478     \__tag_struct_prop_gput:nne{ ####1 }{C}{\l__tbl_tmpa_str}
479   }
480   {
481     \__tag_struct_prop_gput:nne{ ####1 }{C}{[/TH-col~\l__tbl_tmpa_str]}
482   }
483 }
484 {\__tag_struct_prop_gput:nnn{ ####1 }{C}{/TH-col}}
485 }
486 }
487 }
488 }

```

(End of definition for `__tbl_set_header_rows:`)

And some key support: (See `tagpdfsetup-keys.md` for the naming scheme.)

```

489 \keys_define:nn { __tag / setup }
490 {
491   table/header-rows .clist_set:N = \l__tbl_header_rows_clist,
492   table/header-columns .clist_set:N = \l__tbl_header_columns_clist,

```

obsolete older name:

```

493   table-header-rows .meta:n = {table/header-rows={#1}}
494 }

```

5.5 Multirow support

`__tbl_multirow:n` This command makes the current cell into a multirow cell: it creates, if needed, an `RowSpan`-attribute, adds it to the attributes of the cell structure, and marks all following cells spanned by the multirow as cells that should not be tagged. The argument is the number of spanned row (including the current row).

```

495 \cs_if_exist:NT \__tag_struct_prop_gput:nnn
496 {
497   \cs_generate_variant:Nn \__tag_struct_prop_gput:nnn {one}
498 }
499 \cs_new_protected:Npn \__tbl_multirow:n #1
500 {

```

Create an attribute if needed:

```

501   \prop_get:NeNF \g__tag_attr_entries_prop
502     {rowspan-\int_eval:n{#1}}
503     \l__tbl_tmpa_tl

```

```

504 {
505   \__tag_attr_new_entry:ee
506   {rowspan-\int_eval:n{#1}}
507   {/O /Table /RowSpan~\int_eval:n{#1}}
508 }

```

ensure that the attribute is marked as used:

```

509 \seq_gput_left:Ne\g__tag_attr_class_used_seq
510 {\pdf_name_from_unicode_e:n{rowspan-\int_eval:n{#1}}}

```

Get the structure number of the current cell

```

511 \seq_get_right:NN\g__tbl_struct_cur_seq \l__tbl_tmpb_tl

```

If we are in a multicolumn the number can be negative and this must be changed

```

512 \tl_set:Ne \l__tbl_tmpb_tl { \int_abs:n{\l__tbl_tmpb_tl} }

```

now we must update an existing attribute. TODO: simplify this ... (see also colspan handling).

```

513 \prop_get:cnNTF
514 { g__tag_struct_\l__tbl_tmpb_tl _prop }
515 { C }
516 \l__tbl_tmpa_tl
517 {
518   \tl_remove_once:Nn \l__tbl_tmpa_tl {}
519   \tl_remove_once:Nn \l__tbl_tmpa_tl {}
520   \__tag_struct_prop_gput:onef \l__tbl_tmpb_tl }
521   {C}
522   {[ /rowspan-\int_eval:n{#1}~\l__tbl_tmpa_tl]}
523 }
524 {
525   \__tag_struct_prop_gput:onef \l__tbl_tmpb_tl }
526   {C}
527   {[ /rowspan-\int_eval:n{#1}]}
528 }

```

Now mark the spanned cells that should be ignored.

```

529 \__tbl_gset_untagged_row_cells:nn {#1-1}{\g__tbl_span_tl}
530 }

```

(End of definition for __tbl_multirow:n.)

```

531 \keys_define:nn{ __tag / setup }
532 { table/multirow .code:n = {\__tbl_multirow:n {#1} }
533   ,table/multirow .default:n = 1
534 }

```

__tbl_gset_untagged_row_cells:nn

This command stores the row and column numbers of the cells in the following row(s) that should not be tagged as they are a part of a rowspan.

```

535 \cs_new_protected:Npn \__tbl_gset_untagged_row_cells:nn #1 #2
536 % #1 number of rows, #2 number of columns
537 {
538   \int_step_inline:nn {#1}
539   {
540     \int_step_inline:nn {#2}
541     {

```

```

542         \prop_gput:Nee \g__tbl_untagged_cells_prop
543         {
544             \int_eval:n {\g__tbl_row_int + ##1},
545             \int_eval:n{\g__tbl_col_int + #####1 -1 }
546         }{}
547     }
548 }
549 }

```

(End of definition for `__tbl_gset_untagged_row_cells:nn`.)

`__tbl_if_tag_cell:nn` We must be able to detect if a cell should be tagged. For this we define a conditional — if more options are needed it can be extended.

```

550 \prg_new_protected_conditional:Npnn\__tbl_if_tag_cell:nn #1 #2 %#1 row, #2 col
551 { T,TF }
552 {
553     \prop_get:NenTF \g__tbl_untagged_cells_prop
554     {\int_eval:n{##1},\int_eval:n{##2}}\l__tbl_tmpa_tl
555     { \prg_return_false:}
556     { \prg_return_true: }
557 }

```

(End of definition for `__tbl_if_tag_cell:nn`.)

5.6 Misc stuff

`__tbl_show_curr_cell_data:` Show the row/column index and span count for current table cell for debugging.

```

558 \cs_new_protected:Npn \__tbl_show_curr_cell_data: {
559     \__tbl_trace:n { ==>~ current~cell~data:~
560         \int_use:N \g__tbl_row_int ,
561         \int_use:N \g__tbl_col_int ,
562         \g__tbl_span_tl
563     }
564 }

```

(End of definition for `__tbl_show_curr_cell_data:.`)

`__tbl_add_missing_cells:` The storing and use of the number of missing cells must happen at different places as the testing happens at the end of the last cell of a row, but still inside that cell, so we use two commands. The one adding is used in the row/end socket.

```

565 \cs_new:Npn \__tbl_add_missing_cells:
566 {

```

The TD-socket messages are issued after the message about the end-row socket, but the structure is ok, so better issue a message for now to avoid confusion:

```

567     \int_compare:nNnT \g__tbl_missing_cells_int > 0
568     {
569         \__tbl_trace:n {==>~
570             ~Inserting~\int_use:N \g__tbl_missing_cells_int \space
571             additional~cell(s)~into~previous~row:}
572         \int_step_inline:nn { \g__tbl_missing_cells_int }
573         {
574             \int_gincr:N\g__tbl_col_int
575             \UseTaggingSocket{tbl/cell/begin}

```

```

576             \UseTaggingSocket{tbl/cell/end}
577         }
578     }
579 }

```

(End of definition for _tbl_add_missing_cells:.)

\g__tbl_struct_table_tl We need to store the structure numbers for the fine tuning in the finalize socket. For now we use a rather simple system: A sequence that hold the numbers for the row structures, and one that holds comma lists for the cells.

\l__tbl_saved_struct_table_tl \g__tbl_struct_table_tl will hold the structure number of the table, \g__tbl_struct_rows_seq will hold at index i the structure number of row i, \g__tbl_struct_cells_seq will hold at index i a comma list of the cell structure numbers of row i.

\l__tbl_saved_struct_rows_seq \g__tbl_struct_cur_seq is used as a temporary store for the cell structures of the current row. We need also local version to store and restore the values.

\l__tbl_saved_struct_cells_seq

```

580 \tl_new:N \g__tbl_struct_table_tl
581 \tl_new:N \l__tbl_saved_struct_table_tl
582 \seq_new:N \g__tbl_struct_rows_seq
583 \seq_new:N \l__tbl_saved_struct_rows_seq
584 \seq_new:N \g__tbl_struct_cells_seq
585 \seq_new:N \l__tbl_saved_struct_cells_seq
586 \seq_new:N \g__tbl_struct_cur_seq
587 \seq_new:N \l__tbl_saved_struct_cur_seq

```

(End of definition for \g__tbl_struct_table_tl and others.)

_tbl_init_struct_data: Save the global structure data variables locally so the we can restore them when the table ends (important in case of nested tables). This is also the right point to initialize them. \g__tbl_struct_table_tl is set elsewhere.

```

588 \cs_new_protected:Npn \_tbl_init_struct_data: {
589   \tl_set_eq:NN \l__tbl_saved_struct_table_tl \g__tbl_struct_table_tl
590   \seq_set_eq:NN \l__tbl_saved_struct_rows_seq \g__tbl_struct_rows_seq
591   \seq_set_eq:NN \l__tbl_saved_struct_cells_seq \g__tbl_struct_cells_seq
592   \seq_set_eq:NN \l__tbl_saved_struct_cur_seq \g__tbl_struct_cur_seq
593   %
594   \seq_gclear:N\g__tbl_struct_rows_seq
595   \seq_gclear:N\g__tbl_struct_cells_seq
596   \seq_gclear:N\g__tbl_struct_cur_seq
597 }

```

(End of definition for _tbl_init_struct_data:.)

_tbl_restore_struct_data:

```

598 \cs_new_protected:Npn \_tbl_restore_struct_data: {
599   \tl_gset_eq:NN \g__tbl_struct_table_tl \l__tbl_saved_struct_table_tl
600   \seq_gset_eq:NN \g__tbl_struct_rows_seq \l__tbl_saved_struct_rows_seq
601   \seq_gset_eq:NN \g__tbl_struct_cells_seq \l__tbl_saved_struct_cells_seq
602   \seq_gset_eq:NN \g__tbl_struct_cur_seq \l__tbl_saved_struct_cur_seq
603 }

```

(End of definition for _tbl_restore_struct_data:.)

5.7 longtable

`__tbl_patch_LT@makecaption` This patch is quite similar to the one for LaTeX's `\@makecaption` we also have to change the parbox sockets.

```

604 \def\__tbl_patch_LT@makecaption#1#2#3{%
605   \LT@mcol\LT@cols c{%
606     \AssignTaggingSocketPlug{parbox/before}{noop}
607     \AssignTaggingSocketPlug{parbox/after}{noop}
608     \hbox to\z@{\hss\parbox[t]{\LTcapwidth{%
609       \reset@font
610       \tag_suspend:n{caption}
611       \sbox\@tempboxa{#1{#2:~}#3}%
612       \tag_resume:n{caption}
613       \ifdim\wd\@tempboxa>\hsize
614         #1{#2:~}#3%
615       \else
616         \hbox to\hsize{\hfil#1{#2:~}#3\hfil}%
617       \fi
618       \endgraf\vskip\baselineskip}%
619     \hss}}}
```

(End of definition for `__tbl_patch_LT@makecaption`.)

Overwrite the longtable definition. That will probably break somewhere as they are various package which patch too.

```

620 \AddToHook{package/longtable/after}
621 {
622   \cs_set_eq:NN \LT@makecaption\__tbl_patch_LT@makecaption
623 }
624 \</package>
```