langsci-avm

Felix Kopecky*

Version 0.3.0 – February 21, 2023

1 Introduction

langsci-avm allows typesetting of feature structures, or *attribute-value matrices* (AVM), for use in linguistics. The package provides a minimal and easy to read syntax. The package serves the same purpose as Christopher Manning's avm package, but shares no code base with that package. There is a conversion guide in Section 4.6.

To start using langsci-avm, place \usepackage{langsci-avm} in your preamble.

This documentation is structured as follows: Section 2 describes the input syntax for AVMs and their parts. Ways to customise your AVM's layout follow in Section 3, and selected usage cases are presented in Section 4. There's also an administrative and TeXnical appendix at the end of this document, in case you are interested.

1.1 Example

1.2 Acknowledgements

Thanks to Phelype Oleinik for help on recursion and expansion with LATEX3. Thanks to Ahmet Bilal Özdemir and Stefan Müller for their contributions in planning and testing this package.

 $^{{\}tt *mailto:felix.kopecky@langsci-press.org.} \label{thm:press.org} Please \ submit \ bug \ reports \ and \ feature \ requests \ to \ https://github.com/langsci/langsci-avm/issues.$

2 Structuring AVMs

```
\avm \avm [\langle options \rangle] \{\langle structure \rangle\}
```

This root command of the package type sets AVMs in the documen. In the $\{\langle structure \rangle\}$, delimiter characters are processed to open and close (sub-)structures, as described in Section 2.1. Special elements (e.g. tags, operators, type descriptors) are described in Section 2.2. For a description of the layout $\langle options \rangle$, see Section 3.

A (structure) is basically the content of a stylised tabular: The columns are separated by & and a new line is entered with \\.

2.1Entering (sub-)structures within \avm

```
[...]
                   [ \( structure \) ]
<...>
                   < \(structure\) >
(...)
                   ( \langle structure \rangle )
\{...\}
                   \[...\]
                   \[ \langle structure \rangle \]
```

Updated: 2020-10-02 Within the scope of \avm, these delimiters create (sub-)structures that are enclosed by the respective delimiter. Due to the special meaning that curly braces have in LATEX, these are the only ones that need to be run with an escape token (\). It is currently possible to mix delimiters, e.g. with $\langle structure \rangle$), but this may change in future versions.

> langsci-avm expects your (sub-)structures to have at most two columns, so that for every line in each (sub-)structure, there should be no more than one &. It is recommended to have at least some lines with a & in your (structure). Currently, display issues may appear in some structures if none are given – see the align=false option to remedy this effect.

```
\avm{
       [ < ( \{ ... \} ) > ]
                                                                                         [\langle (\{\ldots\}) \rangle]
\sum {avm{}
    [\{ ... \} \\
< ( ... ), ( ... ) > ]
                                                                                        \begin{bmatrix} \{\ldots\} \\ \langle (\ldots), (\ldots) \rangle \end{bmatrix}
```

\[... \]

New: 2020-10-02 Add a semantic bracket $[\langle structure \rangle]$.

Warning: Semantic brackets are only available when the package option [lfg] is loaded (\usepackage[lfg]{langsci-avm}). Documents with this option can only be compiled with X¬ $\exists E$ T¬EX. If the [1fg] option is not present, \[{ $\langle structure \rangle$ } \] will result in none delimiter output, but the $\{\langle structure \rangle\}$ will be printed nonetheless. (The semantic delimiters are not available in every font, and are currently not provided in standard LATEX documents. If you load the [lfg] option but do not provide the symbol (e.g. by using a font such as libertinus), the package unicode-math will automatically be loaded to provide the symbol.)

\lframe ... \rframe \lframe \structure \rframe

New: 2021-03-03 Delimit a (structure) placed in a rectangular box, which is used in Fillmore & Kay's notation. It can be used like the other delimiters.

```
\avm{
  \lframe ... \rframe
}
```

The parameters of the frame can be adjusted with these options:

framewidth = $\langle length \rangle$ (initially 1pt) Width of the frame.

 $framesep = \langle length \rangle$ (initially 3pt)

Separation of the frame and its contents.

!...! ! $\langle text \rangle$!

Escapes the avm mode so that all delimiters can be used as usual characters. If you need ! as a regular character, see Section 3 for how to change the switch.

Commands for tags, types, unusal lines, and relations

\tag $\t {\langle identifier \rangle}$ \0 \0, \1, \2, \3, \4, \5, \6, \7, \8, \9 \1

 $\t y$ puts its ${\langle identifier \rangle}$ in a box, more precisely an $\t y$. Within the box, the tags font is applied. $\0$, $\1$, ..., $\9$ are shortcuts to \t and place the respective number in the box. For example, \4 is equivalent to \tag{4}. The shortcuts do not take any arguments.

If you want to use this command outside an AVM, you can obtain, for example, [4], by using $\wd{4}$, or the equivalent {\fboxsep.25ex\fbox{\footnotesize 4}}.

```
\avm{[ attr1 & \4\\
       attr2 & \4[attr3 & val3\\
                  attr4 & val4] ]}
```

$$\begin{bmatrix} \text{ATTR1 } \boxed{4} \\ \text{ATTR2 } \boxed{4} \begin{bmatrix} \text{ATTR3 } val3 \\ \text{ATTR4 } val4 \end{bmatrix} \end{bmatrix}$$

\type \type*

\9

 $\type(*) \{(type)\}$

Updated: 2020-04-29

Will output the $\langle type \rangle$ in the types font (serif italics by default). The starred variant Updated: 2020-03-30 \type* will span the complete (sub-)structure and can only be placed in the first column of this structure. After the starred \type*, a \\ is recommended, but can usually be omitted.

> \avm{[\type*{A type spanning a line} attr & [\type{type}]]}

 $\begin{bmatrix} A & type & spanning & a & line \\ \text{ATTR} & [type] \end{bmatrix}$

\id

\id $\{\langle id \rangle\}\ \{\langle structure \rangle\}$

New: 2020-10-02 A variant of \substack from amsmath, this command adds an identifier to the Updated: 2021-12-14 $\{\langle structure \rangle\}$. The contents of $\{\langle id \rangle\}$ will be set in math mode by default, which is convenient given that they often contain variables with subscript indices. Multiple IDs should be separated by a new line, $\setminus \setminus$.

```
\avm{\left(n_1\right)_2}{\left[subj\right]_pred\&swim]}}
```

 $\begin{bmatrix}
\text{SUBJ} \\
n_1 \\
n_2
\end{bmatrix}$ PRED swim

The position of the $\{\langle id \rangle\}$ column relative to the $\{\langle structure \rangle\}$ and the alignment within the $\{\langle id \rangle\}$ column can be changed:

id align = $\langle token \rangle$

Change the alignment of the column inserted by \id. Has to be a column specification. The most probable choices are 1 and r.

id position = $\langle option \rangle$

(initially south-west)

Change the position of \id. In the standard setting south-west, the \id is placed in the lower left corner of the enclosed structure. When set to south-east, the contents are set to the lower right corner. Currently, only south-west and south-east are recognised inputs, and an error is raised when an unknown option is input.

\punk

\punk $\{\langle attribute \rangle\}\{\langle type \rangle\}$

Updated: 2021-06-08 Some $\langle attributes \rangle$ think that the layout of the other attributes in their community leaves no space for them to express their individuality. They desire a life outside the confines of the alignment defined by the others, while still remaining a member of the matrix.

> Technically, this is a line with no snapping to the column layout, but with spacing between the $\langle attribute \rangle$ and $\langle type \rangle$. After \punk, a \\ is obligatory if not in the last line.

```
\avm{[ attr1 & val1\\
      \punk{a quite long attr2}{val2}
       attr3 & val3\\
       attr4 & val4
     ]}
```

A QUITE LONG ATTR2 val2
ATTR3 val3
ATTR4 val4

Hint: Also have a look at the option align=false.

\+

Updated: 2020-03-16

In the scope of \avm, \+ comes out as "\theta". "+" can be obtained normally. In the earlier Version 0.1.0-beta, + produced " \oplus ".

In the scope of \avm, \- comes out as "⊖". To use the "optional hyphenation" meaning New: 2020-03-17 of \-, please write !\-!, where ! is your current switch token.

\shuffle

In the scope of \avm, \shuffle is a shortcut for "O" to mark the shuffle relation.

New: 2020-03-17

3 AVM layout

3.1Defining styles

You can customise many aspects of how an AVM is printed, including the fonts or spacing between delimiters and content. You can apply them locally via the $[\langle options \rangle]$ of α or globally by using \avmsetup. And you can also define your own styles and use them via the $[\langle style = \rangle]$ option in \avm.

 $\langle \text{avmsetup } \langle \text{options} \rangle$

{\langle options \rangle} is a comma-separated list of key=value settings. See the list below for all user-configurable options. The $\{\langle options \rangle\}$ are the same as in $\langle options \rangle$. When inserted in \avm[\(\langle options \rangle \], they apply locally, and globally if given to \avmsetup. Local settings always override global ones, and you can have any feasible number of \avmsetups in your document. The scope of \avmsetup can be restricted through grouping.

\avmdefinestyle \avmdefinestyle $\{\langle name \rangle\}$ $\{\langle settings \rangle\}$

New: 2020-05-11 Instead of applying settings globally or per AVM, you can also define styles and assign them to AVMs, as in $\ensuremath{\mathtt{Avm}}[\mathtt{style} = \langle name \rangle] \{...\}$. The $\langle settings \rangle$ are a comma-separated list of key=value settings, and should be a subset of the settings from \avmsetup. For example, the following plain style highlights neither attributes, values, nor types:

> \avmdefinestyle{plain}{attributes=\normalfont, values=\normalfont, types=\normalfont}

The style is applied with \avm[style=plain]{...}.

Now to the list of settings:

 $style = \langle name \rangle$

In addition to any style that you possibly define yourself, a style narrow is predefined in the package (see Section 4.1).

(initially empty)

 $align = \langle choice \rangle$ (initially true)

Controls whether the columns in the AVM and its substructures should be aligned (snapping to the grid) or not. Aligned AVMs are separated by columnsep, nonaligned are separated by vectorsep.

 $stretch = \langle factor \rangle$ (initially 0.9)

Define \arraystretch, i.e. a factor in the determination of line height.

 $columnsep = \langle length \rangle$ (initially 0.5ex)

Define the \tabcolsep, i.e. horizontal space between columns. The first and second column will have O\columnsep to the left and right, respectively. Between the two the distance is 2\columnsep. Using relative units (like ex or em) may be a good idea so that columnsep scales well with changes in font size.

 $vectorsep = \langle length \rangle$ (initially 1em)

Define the horizontal separation between columns in non-aligned matrices (see option align).

 $delimfactor = \langle factor \rangle$

(initially 1000)

Sets \delimiterfactor. The calculation for the minimum height of a delimiter is $y \cdot f/1000$, where y is the height of the content and f the value of delimfactor. The default 1000 ensure that the delimiters' height is at least that of the structure.

 $delimfall = \langle length \rangle$

(initially Opt)

Controls \delimitershortfall, i.e. the maximum height that the delimiters can be shorter than the enclosed structure. The default Opt ensure that the delimiters are not shorter than the contents.

 $extraskip = \langle length \rangle$

(initially \smallskipamount)

If a substructure is immediately followed by a \\, an extra amount of vertical skip is added so that the content of the next line, possibly another delimiter, does not clash with the delimiter in that line. This automatic skip insertion can be circumvented with placing a \relax before the linebreak, i.e. \relax\\.

attributes = $\langle font \ settings \rangle$

(initially \scshape)

The font for attributes, i.e. the first column of each structure.

values = $\langle font \ settings \rangle$

(initially \itshape)

The font for values, i.e. the second column of each structure.

types = $\langle font \ settings \rangle$

(initially \itshape)

The font used in \type and \type*.

 $tags = \langle format \ settings \rangle$

 $(initially \ \verb|\footnotesize|)$

The font (size) used in \t ag and the shortcuts 1...9.

 $\mathtt{switch} = \langle token \rangle$

(initially!)

Define the escape token. Change this if you need to use "!" as a text glyph.

customise = $\langle settings \rangle$

(initially empty)

An interface to input custom commands to be run at the beginning of every \avm.

3.2 Drawing edges between AVM contents

It is possible to make AVM contents available to tikz, so that they can be referenced in a tikzpicture. To enable this feature, langsci-avm has to be loaded with the option [tikz]:

\usepackage[tikz]{langsci-avm}

Additionaly, avm environments on which tikz is to be used need to have the [pic] option present:

\avm[pic] {...}

Only the parts of an AVM that are specifically marked will be known to tikz. To mark a part of an AVM to be used by TikZ, use \node :

```
\node
```

```
\node \{\langle id \rangle\}\ \{\langle contents \rangle\}
```

New: 2020-09-23 $\{\langle id \rangle\}$ serves as part of the node's identifier in a tikzpicture. It will be prefixed, and it's complete name will be avm-n- $\langle id \rangle$, where n is the counter of \avm in your document that have the [pic] option enabled and that don't have a picname (see below). n starts at 1. For example, a \node named "pretty-node" in the fourth [pic]-enabled avm in your document will be avm-4-pretty-node. Note that \node will register the complete name globally in your document, and so can't be declared by other tikz nodes.

> This behaviour can be adjusted by passing a [picname = $\langle avm's name \rangle$] to \avm. E.g., \nodes within \avm[pic, picname=example1] will have a full name pattern of example $1-\langle id \rangle$. Named \avms do not raise the n mentioned in the last paragraph.

> Any (sub-)structure can be placed into $\{\langle contents \rangle\}$. It could be just a value, an attribute's name, or parts thereof, but whole (sub-)structures can be part of $\{\langle contents \rangle\}$

> A tikzpicture with options [remember picture, overlay] enabled can reference langsci-avm's \nodes. This way, TikZ' extensive drawing abilities are available for the decoration of AVMs. Here's a very simple example document:

$$\begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \left\langle \text{GAP-SS} \right\rangle \\ & & & \\ & & & \\ \text{PHON} & \left\langle \boxed{1} \oplus \boxed{2} \right\rangle \end{bmatrix}$$

```
\documentclass{article}
\usepackage[tikz]{langsci-avm}
\usepackage{tikz} % optional, since langsci-avm will load tikz if option
                  % tikz is present
\usetikzlibrary{arrows,arrows.meta}
\avm[pic]{[ subj & \1\\
        comps & <\node{gap}{gap-ss}> \bigskip\\
        \node{phon}{phon} &
                             <\1 \+ \2>
\begin{tikzpicture}[remember picture, overlay]
        \path[{Stealth[]}-{Stealth[]},gray,dashed,in=90,out=270]
        (avm-1-gap.south) edge (avm-1-phon.north);
\end{tikzpicture}
```

Defining input patterns

```
\arrowvert avmdefine command {\langle name \rangle} [\langle label \rangle] {\langle settings \rangle}
```

New: 2020-06-29 Structures often follow specific patterns. For example, AVMs often have a PHON attribute, which is mapped to a list, the entries of which are in italics. \avmdefinecommand can account for this and other input patterns. For example,

```
\avmdefinecommand{custom}{...}
```

will create a command \custom available only in the scope of \avm (this means that you can have a different meaning in the rest of your document). The (settings) will then be applied to the scope in which \custom is called. If an optional $\langle label \rangle$ is given, the label will be printed, in the current font, before the $\langle settings \rangle$ are applied.

\custom generated in this way automatically advances to the value column after the \langle label \rangle is printed. This means that commands generated with \avmdefinecommand should be called in the attribute column of an existing structure. This behaviour can be circumvented with the starred variant \name*, which is automatically generated by \avmdefinecommand as well. However, it seems advisable to use the starred variants

Here's an example for the aforementioned phon pattern:

```
\avmdefinecommand{phon}[phon]
 {
   attributes = \itshape,
   delimfactor = 900,
   delimfall = 10pt
 }
```

This creates a command \phon (and the variant \phon*) within the scope of any \avm. It will print the label phon in the current font and then apply three settings locally: italics for the attribute (first) column, and two settings for very narrow delimiter fitting. This results in:

```
\avm{
     [\type*{word}
                                                                                    \begin{bmatrix} word \\ \text{PHON} & \left\langle lin'gwistiks \right\rangle \\ \text{SYNSEM} & [\ldots] \end{bmatrix}
      \phon <lin'gwistiks>\\
      synsem & [ ... ]
    1
}
```

Note that any other structure type would have worked instead of $\langle \rangle$. But $\langle \rangle$ and any other markers for sub-structures are left unchanged by \phon and other custom commands. This is why the attribute font is changed by \phon, although lin'qwistiks is technically a value. Remember that < creates a new list sub-substructure, and the first content is printed in its attribute font.

4 Applications

4.1 Spacing and size of delimiters

langsci-avm automatically detects if the end of a sub-structure is followed by a line break. This is useful to find cases in which two sub-structures are printed immediately below each other, and to add extra spacing (the extraskip from the options). This automatic detection can be suppressed with \relax. See below for the effect of that detection:

```
\avm{[ [attr1 & val1 \\ attr2 & val2 ] \\ [attr1 & val1 \\ attr2 & val2 ] \\ [attr1 & val1 \\ attr2 & val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 ] \\ [ATTR1 \ val1 \\ ATTR2 \ val2 \]
```

If many delimiters are nested, this occasionally results in larger delimiter sizes. There is a pre-defined narrow style that resets delimfall (to 5pt) and delimfactor (to 997), which are the values recommended in the *TEXbook*. This results in a more compact appearance:

4.2 Disjunctions and other relations

Sometimes AMVs are placed beside other content to express disjunctions or other relations. In langsci-avm this is done naturally:

4.3 Use as a vector

It's possible to use langsci-avm for feature vectors rather than matrices, as may be useful in generative grammar.

4.4 Combinations with gb4e, expex, and linguex

This package works fine with gb4e and its fork langsci-gb4e. To align the example number at the top of your structure, please use \attop from gb4e:

```
\begin{exe}
  \ex\attop{
  \avm{[ attr1 & val1\\
            attr2 & val2\\
            attr3 & val3]}
  \end{exe}
(1) [ATTR1 val1
ATTR2 val2
ATTR3 val3]
```

The same can be achieved with expex using \envup from lingmacros (see below) or using this experimental syntax:

```
\ex \vtop{\strut\vskip-\baselineskip{
    \avm{[ attr1 & val1\\
        attr2 & val2\\
        attr3 & val3]}
}}
```

Examples typed with linguex can be combined with \evnup from lingmacros to align AVMs (many thanks to Jamie Findlay for pointing this out):

4.5 Combinations with forest

This package also works fine with forest. As per the forest documentation, it is recommended to protect any \avm-statements with {} in nodes:

It may happen that extensive AVMs protrude into the space reserved for other forest nodes or edges. In this case, the forest setting for children = {anchor=north} may be useful: (If you like, try this tree without that setting.)

```
\begin{forest}
                                                       Α
  [A, for children = {anchor=north}
     [B] [{\avm{[attr1 & val1\\
     attr2 & a long value val2\\
                                               В
                                                    ATTR1 val1
     attr3 & val3\\
                                                    ATTR2 a long value val2
     attr4 & val4\\
                                                    ATTR3 val3
     attr5 & val5]}} ]
                                                    ATTR4 val4
 ٦
                                                    ATTR5 val5
\end{forest}
```

4.6 Switching from Christopher Manning's avm package

Switching from avm to langsci-avm will require some, though hopefully minimal, changes to the code. In particular, langsci-avm doesn't distinguish between "active" and "passive" modes, there is now a single way of sorting (see \type, which replaces \asort and \osort), and tags are now produced without @ (\4 instead of @4, etc.).

Paths can be printed with a normal |, and \oplus and other relation symbols can be input more easily (see Section 2.1), though the package will also work with | and \odot .

5 Caveats and planned features

1. There are currently no error messages. If you do not receive the intended output, please make sure that your code fits the syntax described in this documentation. If your code is fine but the output is not, please submit a bug report or feature request at https://github.com/langsci/langsci-avm/issues.

These features are planned for the future:

- 2. A check whether the delimiters are balanced, i.e. whether all (sub-)structures are closed by a], }, etc.
- 3. Improve the appearance of (very) large angle brackets so that they vertically span the complete structure they enclose, maybe using scalerel.

6 Implementation

```
1 \*package\
2 \( \mathref{QC=avm} \)
3 \RequirePackage{xparse}[2022/03/26]
4 \RequirePackage{array}
5 \ProvidesExplPackage {langsci-avm}
6 {2023-02-20} {0.3.0}
7 {AVMs and feature structures}
8
9 \msg_new:nnnn {avm} {lfgoptionmissing}
10 { Missing~package~option~lfg~at~line~\msg_line_number: }
```

```
{
       You~issued~a~command~in~line~\msg_line_number:~that~is~only~available~when~
       the~lfg~package~option~is~enabled.
 14
   \msg_new:nnnn {avm} {idpositionunknown}
        Unkown~value~for~option~id~position`~near~line~\msg_line_number:. }
 17
        You~specified~an~unknown~value~for~option~`id~position`.~The~content~of~
 18
        the~id~could~not~be~output.~Please~see~the~manual~for~a~list~of~valid~
 19
        settings.
 20
     }
 21
Let's first check for package options.
 22 \bool_new:N \l__avm_lfg_bool
 23 \bool_new:N \l__avm_tikz_bool
 24 \DeclareOption{tikz}{ \bool_set_true:N \l__avm_tikz_bool }
 25 \DeclareOption{lfg}{ \bool_set_true:N \l__avm_lfg_bool }
 26 \ProcessOptions\relax
Handling for the TikZ package option.
 27 \bool_if:NT \l__avm_tikz_bool
     {
 28
       \RequirePackage{tikz}
 29
       \newcounter{l__avm_picture_counter}
 30
       \tl_new:N \l__avm_picture_name_prefix_tl
 31
```

Handling for the LFG package option: If the semantic bracket is not available at the end of the preamble (i.e.) it was not loaded by another package, load unicode-math to provide the symbol.

This document command initialises an AVM. The first, optional argumet is a key-value list of settings (see \keys_define:nn below) and the second is the AVM itself, given in the syntax described in this documentation.

\avm enters a group so that keys- and macro-assignemts remain local. It then initialises the commands and shortcuts and any user customisation, sets its mode to true and assigns the keys as given in the optional argument (if any). After the parser __avm_-parse:n is called, the group is closed.

```
41 \NewDocumentCommand{\avm}{ O{} +m }
42
    {
      \c_group_begin_token
43
      \keys_set:nn { avm } { #1 }
44
      \__avm_initialise_document_commands:
45
      \__avm_initialise_custom_commands:
46
      \tl_use:N \l__avm_defined_commands_tl
47
      48
      \__avm_parse:n { #2 }
```

(End definition for \aum. This function is documented on page 2.)

\l_avm_mode_bool
\l_avm_parens_tracker
\l_avm_defined_commands_tl
\l_avm_fillmore_kay_box

We need an auxiliary variable to store the current mode. \l__avm_parens_tracker is a stack for a future check whether the delimiters given to \avm are balanced. \l__avm_-defined_commands_tl is a token list that stores any commands provided by the user via \avmdefinecommand. The box \l__avm_fillmore_kay_box is used as a temporary storage to realise Fillmore & Kay's notation.

```
bool_new:N \l__avm_mode_bool

sa \seq_new:N \l__avm_parens_tracker

tl_new:N \l__avm_defined_commands_tl

box_new:N \l_avm_fillmore_kay_box

tl_new:N \l_avm_parsed_tl

int_new:N \l_avm_mode_switch_character_int
```

(End definition for $\l_avm_mode_bool$ and others.)

\avmsetup

Forward the key-value settings given as the optional argument to \avm to the keys defined in \keys_define:nn { avm }. For the meaning of these keys and initial values, see Section 2.

```
\NewDocumentCommand{\avmsetup}{ m }
    { \keys_set:nn { avm } { #1 } }
59
60
  \keys_define:nn { avm }
      align .bool_set:N
                              = \l_avm_align_bool,
      align .initial:n
                               = {true},
64
                               = \l_avm_arraystretch_tl,
      stretch .tl_set:N
65
      stretch .initial:n
                               = \{0.9\},
66
                              = \l_avm_tabcolsep_dim,
      columnsep .dim_set:N
67
      columnsep .initial:n
                              = \{.5ex\},\
68
      vectorsep .dim_set:N
                              = \l_avm_singlesep_dim,
69
      vectorsep .initial:n
                              = \{1em\},
70
      delimfactor .int_set:N = \l__avm_delimfactor_int,
71
      delimfactor .initial:n = {1000},
      delimfall .dim_set:N = \l__avm_delimshortfall_dim,
      delimfall .initial:n
                              = \{0pt\},\
74
      framewidth .dim_set:N = \l__avm_fillmore_kay_boxrule_dim,
75
      framewidth .initial:n = \{1pt\},
76
      {\tt framesep .dim\_set:} {\tt N}
                               = \l__avm_fillmore_kay_boxsep_dim,
      framesep .initial:n
                               = {3pt},
78
      attributes .code:n
                               = {\cs_set:Nn \__avm_font_attribute: {#1}},
79
      attributes .initial:n = {\scshape},
80
                               = {\cs_set:Nn \__avm_font_type: {#1}},
      types .code:n
81
      types .initial:n
                              = {\itshape},
      values .code:n
                              = {\cs_set:Nn \__avm_font_value: {#1}},
      values .initial:n
                              = {\itshape},
      tags .code:n
                              = {\cs_set:Nn \__avm_font_tag: {#1}},
85
      tags .initial:n
                               = {\footnotesize},
86
      singleton .code:n
                              = {\cs_set:Nn \__avm_font_singleton: {#1}},
87
      singleton .initial:n
                              = {\normalfont},
88
      switch .code:n
89
```

```
90
           \tl_set:Nn \l__avm_mode_switch_character {#1}
91
           \exp_args:NNx \int_set:Nn \l__avm_mode_switch_character_int
92
             {`\tl_use:N \l_avm_mode_switch_character}
93
94
       switch .initial:n
                               = { ! },
95
       extraskip .dim_set:N
                              = \l_avm_extra_skip_dim,
96
       extraskip .initial:n
                              = {\smallskipamount},
       extraskip~in~every~row .bool_set:N = \l__avm_extraskip_bool,
                              = {\cs_set:Nn \__avm_initialise_custom_commands:
       customise .code:n
99
                                             {#1}},
100
       customise .initial:n
                              = { },
101
      pic .bool_set:N
                              = \l__avm_picture_bool,
102
      pic .default:n
                              = { true },
103
      picname .tl_set:N
                              = \l_avm_picture_name_tl,
104
       picname .initial:n
                              = {automatic},
105
                              = { \newcolumntype{i}{#1} },
       id~align .code:n
106
       id~align .initial:n
                              = \{1\},
       id~position .tl_set:N = \l__avm_id_position_tl,
       id~position .initial:n = {south-west},
       style .choice:,
       style / narrow .code:n = {\int_set:Nn \l__avm_delimfactor_int {997}
                               \dim_set:Nn \l__avm_delimshortfall_dim {5pt}},
    }
```

(End definition for \aumsetup. This function is documented on page 5.)

\avmdefinestyle Define a style

Define a style to be used together with the style key.

(End definition for \aumdefinestyle. This function is documented on page 5.)

\avmdefinecommand

A factory function that creates commands for the layout of sub-structures and saves them to \l__avm_defined_commands_tl. The first argument describes the command's name, the second any (optional) label. The manufactured definitions are activated in the AVM group so that they remain local.

(End definition for \aumdefinecommand. This function is documented on page 8.)

\tl_if_eq:VnTF

A useful variant for comparing the values of token list variables with token lists.

```
\cs_generate_variant:Nn \tl_if_eq:nnTF {VnTF}
```

(End definition for \tl_if_eq:VnTF.)

\l_avm_in_first_column

A boolean to check whether we are in the first column (value true) or in the second (value false).

```
132 \bool_new:N \l__avm_in_first_column
```

 $(End\ definition\ for\ \l__aum_in_first_column.)$

_avm_init_first_column:
_avm_init_second_column:
_avm_init_single_column:

These macros apply the settings for the columns in a (sub-)structure. They take care of font selection and report the currently active column back to the system. Knowing which column is active is important when closing the (sub-)structure. If the structure is closed without a second column present, we need to skip back 2\tabcolsep. (This does not apply to the case of vector structures, which are handled without this check.)

```
\cs_new:Nn \__avm_init_first_column:
     {
134
       \bool_set_true:N \l__avm_in_first_column
135
       \normalfont\__avm_font_attribute:
136
137
138
   \cs_new:Nn \__avm_init_second_column:
139
140
        \bool_set_false:N \l__avm_in_first_column
141
       \normalfont\__avm_font_value:
142
     }
143
   \cs_new:Nn \__avm_init_single_column:
145
146
     {
147
       \normalfont\__avm_font_attribute:
     }
148
149
```

_avm_deinit_first_column: _avm_deinit_second_column:

These commands control settings that are applied after each column is exited. The single check here is whether italics is currently in use. If it is, the the italic correction is automatically applied. This replaces the user-configurable setting apptovalues from previous versions.

```
164
                                       {
                                         \tl_if_eq:NNT \f@shape \l__avm_italics_tl {\/}
                                 165
                                 166
                                 (End\ definition\ for\ \verb|\_avm_deinit_first_column:\ and\ \verb|\_avm_deinit_second_column:.|)
\__avm_kern_unused_columns:
                                 A helper macro to fill the horizontal space if a row is ended prematurely, i.e. if no & is
                                 present.
                                    \cs_new:Nn \__avm_kern_unused_columns:
                                 167
                                       {
                                 168
                                         \bool_if:NTF \l__avm_in_first_column
                                 169
                                            { \span\hspace*{-2\tabcolsep} }
                                              }
                                            {
                                 (End\ definition\ for\ \\_avm\_kern\_unused\_columns:.)
                                 This function is used together with the delimiter replacements. It checks whether the
          \__avm_extra_skip:
                                 delimiter is followed by a line break, in which case an extra skip is automatically inserted
                                 173 \cs_new:Nn \__avm_extra_skip:
                                 174
                                         \peek_meaning_ignore_spaces:NTF \\ {\vspace*{\l__avm_extra_skip_dim}} {}
                                 175
                                       }
                                 (End\ definition\ for\ \\_avm\_extra\_skip:.)
          _avm_module_begin:
                                 The replacement instructions for \ \ \  avm_parse:n. When option \ \langle align=true\rangle (default),
                                 the structure has two columns. Vector structures are inserted if \langle align = false \rangle.
          \__avm_module_end:
                          etc.
                                    \cs_new:Nn \__avm_module_begin:
                                 178
                                         \bool_if:NTF \l__avm_align_bool
                                 179
                                 180
                                 181
                                              \begin{tabular}{@{}
                                                                >{\__avm_init_first_column:}l
                                                                <{\__avm_deinit_first_column:}
                                                                >{\__avm_init_second_column:}1
                                                                <{\__avm_deinit_second_column:}
                                 185
                                                                @{}}
                                 186
                                           }
                                 187
                                           {
                                 188
                                              \begin{tabular}{@{}
                                 189
                                                               >{\__avm_init_single_column:}l
                                 190
                                                               <{\__avm_deinit_single_column:}
                                 191
                                                               @{}}
                                 192
                                           }
                                 194
                                      }
                                 195
                                    \cs_new:Nn \__avm_module_end:
                                 196
                                            _avm_kern_unused_columns:
                                 197
                                         \end{tabular}
                                 198
                                 199
                                 200
                                    \cs_new:Nn \__avm_replace_ampersand:
```

\cs_new:Nn __avm_deinit_single_column:

```
202
       \bool_if:NTF \l__avm_align_bool
203
         { \tl_build_put_right: Nn \l__avm_parsed_tl { & } }
204
         { \tl_build_put_right:Nn \l__avm_parsed_tl
205
206
              \exp_not:n
207
208
                  \__avm_deinit_first_column:\skip_horizontal:N
                  \dim_use:N \l__avm_singlesep_dim \__avm_init_second_column:
            }
         }
    }
  \cs_new:Nn \__avm_replace_lbrace:
215
     {
216
       \c_math_toggle_token\left\lbrace\__avm_module_begin:
217
218
  \cs_new:Nn \__avm_replace_rbrace:
219
       \__avm_module_end:\right\rbrace\c_math_toggle_token\__avm_extra_skip:
    }
222
  \cs_new:Nn \__avm_replace_lbrack:
     {
224
       \tl_build_put_right:Nn \l__avm_parsed_tl
225
226
           \exp_not:n
227
228
                \bool_if:NTF \l__avm_mode_bool
                    \c_math_toggle_token\left\lbrack\__avm_module_begin:
                  }
                  { [ }
             }
234
         }
235
    }
236
  \cs_new:Nn \__avm_replace_rbrack:
237
238
239
       \tl_build_put_right:Nn \l__avm_parsed_tl
240
           \exp_not:n
                \bool_if:NTF \l__avm_mode_bool
244
                    \__avm_module_end:\right\rbrack\c_math_toggle_token%
245
                    \__avm_extra_skip:
246
                  }
247
                  { ] }
248
             }
249
         }
250
251
    }
  \bool_if:NTF \l__avm_lfg_bool
253
       \cs_new:Nn \__avm_replace_llbrack:
254
         {
255
```

```
\c_math_toggle_token\left\lBrack\__avm_module_begin:
         }
257
       \cs_new:Nn \__avm_replace_rrbrack:
258
259
             \verb| \_avm_module_end: \verb| right\rBrack\c_math_toggle_token\c_avm_extra_skip: \\
260
261
     }
262
     {
263
       \cs_new:Nn \__avm_replace_llbrack:
           \tl_build_put_right:Nn \l__avm_parsed_tl
              {
267
                \exp_not:n
268
                  {
269
                     \msg_warning:nn {avm}{lfgoptionmissing}
                     \c_math_toggle_token\left.\__avm_module_begin:
              }
273
         }
       \cs_new:Nn \__avm_replace_rrbrack:
275
           \tl_build_put_right:Nn \l__avm_parsed_tl
277
              {
278
                \exp_not:n
279
280
                  \msg_warning:nn {avm}{lfgoptionmissing}
281
                   \__avm_module_end:\right.\c_math_toggle_token\__avm_extra_skip:
282
283
              }
284
         }
     }
   \cs_new:Nn \__avm_replace_lparen:
288
       \tl_build_put_right:Nn \l__avm_parsed_tl
289
290
           \exp_not:n
291
              {
292
                \bool_if:NTF \l__avm_mode_bool
293
294
                    \c_math_toggle_token\left(\__avm_module_begin:
                  }
                  { ( }
              }
298
         }
299
     }
300
   \cs_new:Nn \__avm_replace_rparen:
301
302
       \tl_build_put_right:Nn \l__avm_parsed_tl
303
304
305
            \exp_not:n
                 \bool_if:NTF \l__avm_mode_bool
308
                      \__avm_module_end:\right)\c_math_toggle_token\__avm_extra_skip:
309
```

```
}
{ ) }
310
311
312
313
     }
314
   \cs_new:Nn \__avm_replace_langle:
315
316
       \tl_build_put_right:Nn \l__avm_parsed_tl
317
318
319
            \exp_not:n
              {
                \bool_if:NTF \l__avm_mode_bool
321
                  {
322
                     \c_math_toggle_token\left<\__avm_module_begin:</pre>
323
324
                  { < }
325
              }
326
         }
327
     }
328
   \cs_new:Nn \__avm_replace_rangle:
     {
330
       \tl_build_put_right:Nn \l__avm_parsed_tl
331
332
            \exp_not:n
333
              {
334
                \bool_if:NTF \l__avm_mode_bool
335
336
                  {
                       _avm_module_end:\right>\c_math_toggle_token\__avm_extra_skip:
337
                  }
338
                  { > }
              }
340
         }
341
     }
342
   \cs_new:Nn \__avm_replace_lframe:
343
344
       \hbox_set:Nw \l__avm_fillmore_kay_box \group_begin:
345
       \c_math_toggle_token\__avm_module_begin:
346
     }
347
348
   \cs_new:Nn \__avm_replace_rframe:
       \__avm_module_end:\c_math_toggle_token\group_end:\hbox_set_end:
       \group_begin:
       \dim_set_eq:NN \fboxrule \l__avm_fillmore_kay_boxrule_dim
352
       \dim_set_eq:NN \fboxsep \l__avm_fillmore_kay_boxsep_dim
353
       \fbox{\box_use:N \l__avm_fillmore_kay_box}
354
       \group_end: \__avm_extra_skip:
355
     }
356
   \cs_new:Nn \__avm_replace_plus:
357
358
     {
359
       \leavevmode\unskip\hbox{${}\oplus{}$}\ignorespaces
     }
361
   \cs_new:Nn \__avm_replace_minus:
362
     {
       \verb|\label{leavevmode|unskip|hbox{${}} ominus{}$} ignore spaces
363
```

```
}
                     \cs_new:Nn \__avm_replace_circle:
                366
                              \leavevmode\unskip\hbox{${}\bigcirc{}$}\ignorespaces
               367
               368
              (End definition for \__aum_module_begin:, \__aum_module_end:, and etc..)
 \tag
\type
               369 \cs_new:Npn \__avm_controls_tag:n #1
\punk
                         { \fboxsep.25ex\fboxrule.4pt\fbox{\normalfont\__avm_font_tag: #1} }
\node
                     \cs_new:Npn \__avm_controls_type:n #1
                         { \c_group_begin_token\normalfont\__avm_font_type: #1\c_group_end_token }
               373 \cs_new_protected:Npn \__avm_controls_type_starred:n #1
               374
                              \bool_set_false:N \l__avm_in_first_column
               375
                              \normalfont\__avm_font_type: #1
               376
                              \bool_if:NTF \l__avm_align_bool
                                  { \__avm_deinit_second_column:\span\hspace*{-2\tabcolsep} }
               378
                                  { \_avm_deinit_single_column:}
               379
                              \peek_meaning_ignore_spaces:NTF \\ {} {\\}
               380
                         }
                381
                     \cs_new_protected:Npn \__avm_controls_punk:nn #1 #2
                383
                              \bool_set_false:N \l__avm_in_first_column
                384
                              \normalfont\c_group_begin_token\__avm_font_attribute:#1%
                385
                              \c_group\_end\_token\hspace{2\tabcolsep}{%}% \label{lem:c_group_end_token_hspace}% \label{lem:c_group_end_to
                386
                             \c_group_begin_token\__avm_font_value: #2\c_group_end_token%
                387
                              \__avm_deinit_second_column:\span\hspace*{-2\tabcolsep}
                388
                              \peek_charcode_ignore_spaces:NTF \\ {} {\\}
                389
                390
                391
                      \cs_new:Nn \__avm_mode_switch:
                392
                393
                              \bool_set_inverse:N \l__avm_mode_bool
                              \bool_if:NTF \l__avm_mode_bool
                                      \DeclareDocumentCommand{\{}{}{ \__avm_replace_lbrace: }
                397
                                      \DeclareDocumentCommand{\}}{}{ \__avm_replace_rbrace: }
                398
                                      \DeclareDocumentCommand{\[}{}{ \__avm_replace_llbrack: }
                399
                                      \DeclareDocumentCommand{\]}{}{ \__avm_replace_rrbrack: }
                400
                                      \DeclareDocumentCommand{\+}{}{ \__avm_replace_plus: }
               401
                                      \DeclareDocumentCommand{\-}{}{ \__avm_replace_minus: }
                402
                                  }
                403
                                      \DeclareCommandCopy{\{}{\__avm_old_lbrace_store:}
                                      \DeclareCommandCopy{\}}{\__avm_old_rbrace_store:}
                406
                                      407
                                      \DeclareCommandCopy{\]}{\__avm_old_rrbrack_store:}
               408
                                      \DeclareCommandCopy{\+}{\__avm_old_plus_store:}
               409
                                       \DeclareCommandCopy{\-}{\__avm_old_minus_store:}
               410
               411
                         }
               412
```

413

```
\cs_new:Nn \__avm_initialise_document_commands:
    {
415
      \DeclareCommandCopy{\__avm_old_lbrace_store:}{\{}
416
      \DeclareCommandCopy{\__avm_old_rbrace_store:}{\}}
417
      \DeclareCommandCopy{\__avm_old_llbrack_store:}{\[}
418
      \DeclareCommandCopy{\__avm_old_rrbrack_store:}{\]}
419
      \DeclareCommandCopy{\__avm_old_plus_store:}{\+}
420
      \DeclareCommandCopy{\__avm_old_minus_store:}{\-}
421
      \def\arraystretch{\tl_use:N \l__avm_arraystretch_tl}
422
      \dim_set_eq:NN \tabcolsep \l__avm_tabcolsep_dim
      \int_set_eq:NN \delimiterfactor \l__avm_delimfactor_int
424
      \dim_set_eq:NN \delimitershortfall \l__avm_delimshortfall_dim
425
      \DeclareDocumentCommand{\shuffle}{}{ \__avm_replace_shuffle: }
426
      \DeclareDocumentCommand{\lframe}{}{ \__avm_replace_lframe: }
427
      \DeclareDocumentCommand{\rframe}{}{ \__avm_replace_rframe: }
428
      \DeclareDocumentCommand{\tag}{m}{ \__avm_controls_tag:n {##1} }
429
      \DeclareDocumentCommand{\0}{}{ \__avm_controls_tag:n {0} }
      \DeclareDocumentCommand{\1}{}{ \__avm_controls_tag:n {1}
      \DeclareDocumentCommand{\3}{}{ \__avm_controls_tag:n {3} }
      \DeclareDocumentCommand{\4}{}{ \__avm_controls_tag:n {4} }
      \DeclareDocumentCommand{\5}{}{ \__avm_controls_tag:n {5} }
435
      \DeclareDocumentCommand{\6}{}{ \__avm_controls_tag:n {6} }
436
      437
      \DeclareDocumentCommand{\8}{}{ \__avm_controls_tag:n {8} }
438
      \DeclareDocumentCommand{\9}{}{ \__avm_controls_tag:n {9} }
439
      \DeclareDocumentCommand{\type}{s m}
440
441
          \IfBooleanTF { ##1 }
            { \__avm_controls_type_starred:n {##2} }
111
            { \_avm_controls_type:n {##2} }
445
      \DeclareDocumentCommand{\punk}{m m}{ \__avm_controls_punk:nn {##1}{##2} }
446
      \DeclareDocumentCommand{\id}{m m}
447
448
          \hcoffin_set:Nw \l_tmpa_coffin
            \bgroup
450
            \def\arraystretch{.5}
451
452
            \begin{tabular}[b]{0{}>{$\scriptstyle}i<{$}0{}}
            ##1
            \end{tabular}
            \egroup
            \hcoffin_set_end:
            \hcoffin_set:Nw \l_tmpb_coffin ##2 \hcoffin_set_end:
457
            \tl_if_eq:VnTF \l__avm_id_position_tl {south-west}
458
              ₹%
459
                \coffin_join:NnnNnnn \l_tmpb_coffin {1}{H}
460
                 \l_tmpa_coffin {r}{H}{ Opt }{ -\coffin_dp:N \l_tmpb_coffin }
461
              }
462
                \tl_if_eq:VnTF \l__avm_id_position_tl {south-east}
                {%
                  \coffin_join:NnnNnnnn \l_tmpb_coffin {1}{H}
466
                  \l_tmpa_coffin {1}{H}{ \coffin_wd:N \l_tmpb_coffin }
467
```

```
{ -\coffin_dp:N \l_tmpb_coffin }
468
                  }
469
                  {
470
                     \msg_error:nn {avm}{idpositionunknown}
471
                  }
472
                }
473
            \coffin_typeset:Nnnnn \l_tmpb_coffin {l}{vc}{0pt}{0pt}
474
475
The last of the bunch is only loaded if TikZ is loaded as well:
       \bool_if:NT \l__avm_tikz_bool
            \tl_if_eq:VnTF \l__avm_picture_name_tl {automatic}
479
                \stepcounter{l__avm_picture_counter}
480
                \tl_set:Nn \l__avm_picture_name_prefix_tl
181
                   \{avm-\tl\_use: \t \ \thel\_avm\_picture\_counter\} 
482
              }
483
484
                \tl_set_eq:NN \l__avm_picture_name_prefix_tl \l__avm_picture_name_tl
485
              }
            \DeclareDocumentCommand{\node}{m m}
                \tikz [remember~picture,
                        baseline=(\l__avm_picture_name_prefix_tl-##1.base)]
                \node [inner~sep=0pt] (\l__avm_picture_name_prefix_tl-##1)
                                         {\strut ##2};
492
              }
493
         }
494
     }
495
```

(End definition for \tag and others. These functions are documented on page 3.)

__avm_parse:n F

Finally, the parser. It is build on \@@_act:NNNnn from 13t1 (see the sub-section *Token by token changes*). Many thanks to Phelype Oleinik for help on this, and in particular on help with expansion.

```
\cs_new:Npn \__avm_parse:n #1
496
       \group_align_safe_begin:
       \tl_build_begin:N \l__avm_parsed_tl
499
       \tl_build_put_right:Nn \l__avm_parsed_tl { \exp_not:n {\ignorespaces} }
       \tl_analysis_map_inline:nn { #1 }
501
         {
502
           \int_case:nnF { ##2 }
503
504
               { `& }{ \_avm_replace_ampersand: }
505
               { `[ }{ \_avm_replace_lbrack: }
506
               { `] }{ \_avm_replace_rbrack: }
               { `( }{ \_avm_replace_lparen: }
               { `) }{ \__avm_replace_rparen: }
               { `< }{ \sum_{avm\_replace\_langle: } }
510
               { `> }{ \__avm_replace_rangle: }
511
               { \l_avm_mode_switch_character_int }
512
                 {
513
                   \tl_build_put_right:Nn \l__avm_parsed_tl
514
```

```
{ \exp_not:n { \__avm_mode_switch: } }
515
                    }
516
               }
517
               {
518
                  \tl_build_put_right:Nn \l__avm_parsed_tl { ##1 }
519
520
          }
521
        \tl_build_end:N \l__avm_parsed_tl
522
        \verb|\tl_set:Nx \l__avm_parsed_tl {\l__avm_parsed_tl}|
523
        \tl_use:N \l__avm_parsed_tl
524
        \group_align_safe_end:
525
     }
526
527
(End\ definition\ for\ \\_aum\_parse:n.)
_{528} \langle /package \rangle
```