

BIBULOUS

A drop-in BibTeX replacement based on style templates.

Bibulous Documentation

Release 1.3.2

Bibulous developers

CONTENTS

1	Getting started 1.1 Kile: replacing BibTeX with Bibulous	3
	1.2 Modifying WinEdt5 to replace BibTeX with Bibulous	4
	1.3 Modifying Texmaker to replace BibTeX with Bibulous	5
	1.4 Modifying TeXStudio to replace BibTeX with Bibulous	6
	1.5 Modifying TeXworks to replace BibTeX with Bibulous	7
	1.6 Modifying LyX to replace BibTeX with Bibulous	7
2	Frequently asked questions	9
	2.1 1. What is the difference between using Bibulous and using BibLatex+Biber?	9
3	- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	11
		11
		13
		17 19
		20
	5.5 Tytholi ATT	20
4		25
		25
	4.2 How to report a bug	25
5		27
		27
	1 3 63	27
	e	28
	e	29
	ϵ	29
	e	30 31
		31
		32
	ϵ	32
	5.11 Updating the PyPI package	33
		33
6	Examples	35
0	1	35
	1	36
	1	39
	1	44

6	Example 5
6.0	
6.	Example 7
6.	
6.9	Example 9
6.	0 Example 10
	1 Example 11
7 Bi	pulous Overview
7.	Installation
7.	Example
7.3	Developers
7.4	Contact
7.	Thomas
/	License

Contents:

CONTENTS 1

2 CONTENTS

GETTING STARTED

For general users, all that is needed is place the main bibulous.py file into the Python path and to tell LaTeX to use Bibulous rather than BibTeX as their bibliography engine. For users interested in using the auxiliary command bibulous_authorextract.py, this file must also be in the Python path, and must be in the same directory as the main file. It is also possible to tell any LaTeX front-end to use Bibulous in place of BibTeX. For Kile, Texmaker, TeXStudio, TeXworks, and Lyx (on Linux) and WinEdt (on Windows), instructions for doing this are given below.

Once you bibliography engine is set to Bibulous, remember that the filename within the \bibliography{...} command in your *.tex files must point to a Bibulous-format bibliography style template (.bst file) and not a BibTeX one. Note that if the bibulous.py file is placed in a standard Python path location, then its complete path does not need to be specified in each setup below. For example, if Bibulous is installed using pip install bibulous, then the Python path will be updated appropriately such that any call to bibulous.py will work.

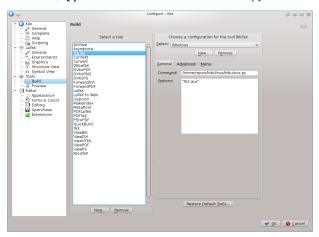
For users who run the bibliography engine at the command line:

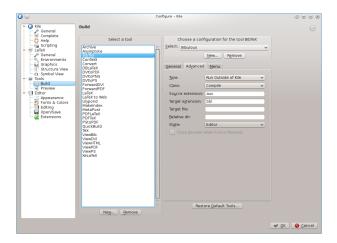
python /path_to_file/bibulous.py mybib.aux

where path_to_file gives the location of the bibulous.py file, and mybib.aux gives the name of the file to process. The output file will have the same filename as the aux file, but a different extension, i.e. mybib.bbl.

1.1 Kile: replacing BibTeX with Bibulous

1. In Kile, go to the menu bar and select Settings > Configure Kile. Select Tools > Build and choose BibTeX from the Select a tool menu (see the figure). To the right of the menu, after you select BibTeX you should see "Choose a configuration for the tool BibTeX". Below the drop-down menu, select the button "New" and type in the name Bibulous (or whatever you prefer to call your new tool). Below, in the General tab, type in the location of the bibulous.py file. And in the Options field, type %dir_base/%S.aux.





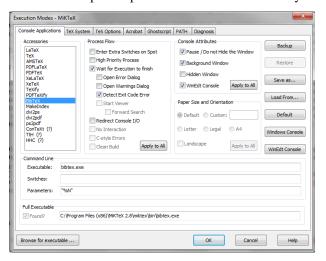
That should be it. In case your default setup is different, you can also check the Advanced tab settings and verify that they are as shown in the second figure. (That is, Source extension is set to aux, and Target extension is set to bbl.)

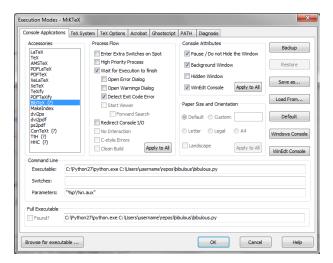
2. Note that the following variables are accessible in Kile's Options field:

```
%source = filename (i.e. filename with suffix but not path)
%S = filename without suffix (and without path)
%dir_base = source file directory (source file's path without the filename)
%dir_target = target file directory (source file's path without the filename)
```

1.2 Modifying WinEdt5 to replace BibTeX with Bibulous

1. Go to the menu Options > Execution Modes. In the Console Applications menu on the left hand side, select BibTeX. Then replace the three Command Line fields with the ones shown in the figure, replacing the example file paths with the ones correct for your installation of Python and bibulous.py.



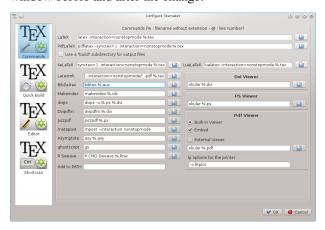


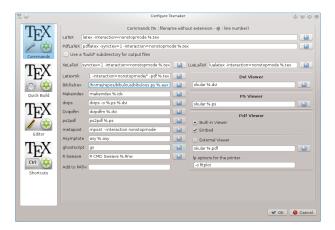
2. Note that the following are definitions of WinEdt registers:

```
%f = full path of active file (= %p/%n.%t)
%n = name of the active file
%p = the path of the active file
%t = the extension of the active file
%q = the path relative to the main file (i.e. for subdirectories)
%b = WinEdt's local working directory (not the tex file directory)
%B = path to the WinEdt executable file
```

1.3 Modifying Texmaker to replace BibTeX with Bibulous

On the main menu bar, select Options then Configure Texmaker. In the Bib (la) tex field, replace bibtex %.aux with the location of your bibulous.py file, followed by %.aux. The following figures show the configuration window before and after the change:

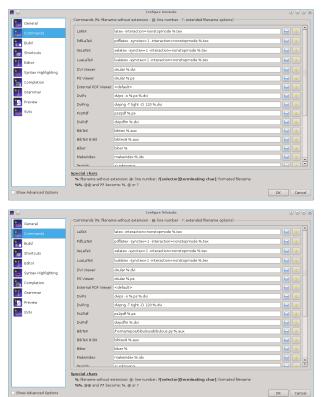




(The example shows the location of the bibulous.py file as /home/repos/bibulous/bibulous.py, but you need to replace it with the appropriate location on your own system.) Also note that the bibulous.py must be executable by the user.

1.4 Modifying TeXStudio to replace BibTeX with Bibulous

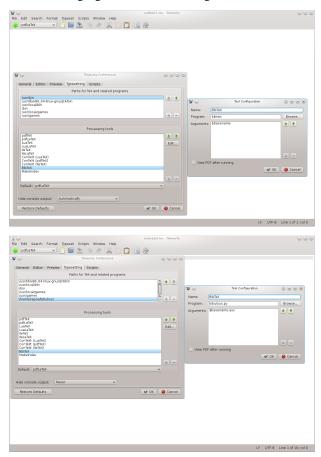
On the main menu bar, select Options then Configure TeXStudio, and then the Commands tab. In the BibTeX field, replace bibtex %.aux with the location of your bibulous.py file, followed by %.aux. The following figures show the configuration window before and after the change:



(The example shows the location of the bibulous.py file as /home/repos/bibulous/bibulous.py, but you need to replace it with the appropriate location on your own system.)

1.5 Modifying TeXworks to replace BibTeX with Bibulous

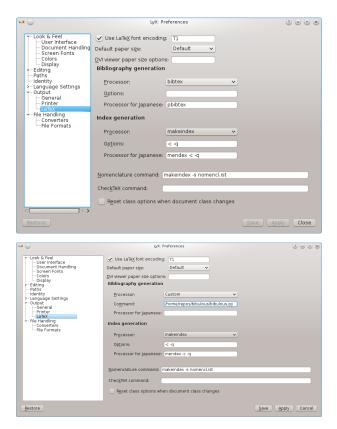
On the main menu bar, select Edit then Preferences..., and then the Typesetting tab. In the Paths for TeX and related programs list, insert a new entry to give the path to your bibulous.py program. (This is the highlighted entry in the figure shown below.) In the Processing tools list, double-click on BibTeX. This brings up a new dialog window. In that window, remove bibtex from the Program field and replace it with bibulous.py. The following figures show the configuration windows before and after the change:



(The example shows the location of the bibulous.py file as /home/repos/bibulous/bibulous.py, but you need to replace it with the appropriate location on your own system.)

1.6 Modifying LyX to replace BibTeX with Bibulous

On the main menu bar, select Tools then Preferences, and then expand the Output dropdown list and click on the LaTeX field. In the Bibliography generation section of the resulting dialog, change Processor to Custom and in the Command field insert the full path to your bibulous.py program. The following figures show the configuration windows before and after the change:



(The example shows the location of the bibulous.py file as /home/repos/bibulous/bibulous.py, but you need to replace it with the appropriate location on your own system.)

FREQUENTLY ASKED QUESTIONS

2.1 1. What is the difference between using Bibulous and using BibLatex+Biber?

The biggest difference between these two is that Bibulous uses bibliography style template files that are intuitive and compact, whereas the Bibulous+Biber combination uses LaTeX code to generate styles, and these make them much more difficult to create and customize. A question posted on StackExchange [http://tex.stackexchange.com/questions/151628/bibtex-fields-for-doi-mr-zbl-and-arxiv/163628#163628] asked how one can build a reference list with hyperlinks to zbl, mr, doi, and eprint fields as defined in the database entries.

For Bibulous, we can create a template file, or modify an existing one, to have a link variable (you can call it almost anything you like), and add a definition for link in the SPECIAL-TEMPLATES section of the file:

The definition for link in the SPECIAL-TEMPLATES section creates a variable that can be used in the entrytype style templates. Thus, any template can use link> to insert a hyperref to the quantities defined in the database. Adding the definition, and inserting <link> in each template where desired is all that needs to be done. The square brackets around <link> tell the template to ignore it if none of the four fields is defined inside the database entry.

For BibLatex+Biber, we can compare to Bibulous by using the answer provided on the same StackExchange page, given as follows. We add the following lines to the preamble of the main tex file:

(continues on next page)

(continued from previous page)

```
\DeclareFieldFormat{arxiv}{%
arXiv\addcolon\space
\ifhyperref
    {\href{http://arxiv.org/\abx@arxivpath/#1}{%
    \nolinkurl{#1}%
    \iffieldundef{arxivclass}
        {\addspace\texttt{\mkbibbrackets{\thefield{arxivclass}}}}}}
    {\nolinkurl{#1}
    \iffieldundef{arxivclass}
    {\addspace\texttt{\mkbibbrackets{\thefield{arxivclass}}}}}}
\makeatother
\DeclareFieldFormat{pmcid}{%
PMCID\addcolon\space
\ifhyperref
    {\href{http://www.ncbi.nlm.nih.gov/pmc/articles/#1}{\nolinkurl{#1}}}
    {\nolinkurl{#1}}}
\DeclareFieldFormat{mr}{%
MR\addcolon\space
\ifhyperref
    {\href{http://www.ams.org/mathscinet-getitem?mr=MR#1}{\nolinkurl{#1}}}
    {\nolinkurl{#1}}}
\DeclareFieldFormat{zbl}{%
Zbl\addcolon\space
\ifhyperref
    {\href{http://zbmath.org/?q=an:#1}{\nolinkurl{#1}}}
    {\nolinkurl{#1}}}
\DeclareFieldAlias{jstor}{eprint:jstor}
\DeclareFieldAlias{hdl}{eprint:hdl}
\DeclareFieldAlias{pubmed} {eprint:pubmed}
\DeclareFieldAlias{googlebooks}{eprint:googlebooks}
\renewbibmacro*{eprint}{%
\printfield{arxiv}%
\newunit\newblock
\printfield{jstor}%
\newunit\newblock
\printfield{mr}%
\newunit\newblock
\printfield{zbl}%
\newunit\newblock
\printfield{hdl}%
\newunit\newblock
\printfield{pubmed}%
\newunit\newblock
\printfield{pmcid}%
\newunit\newblock
\printfield{googlebooks}%
\newunit\newblock
\iffieldundef{eprinttype}
    {\printfield{eprint}}
    {\printfield[eprint:\strfield{eprinttype}]{eprint}}}
```

The template approach is clearly more compact, easier to read, and easier to customize.

GUIDELINES FOR WRITING BIBLIOGRAPHY STYLE TEMPLATES

3.1 Syntax

- 1. Comments begin with #, following the Python convention.
- 2. Each template file can have as many as five sections. None of the sections are required to be in the file, but any definitions in the file must be placed inside a section header so that the code knows how to deal with the definition. The four possible section headers are: TEMPLATES, SPECIAL-TEMPLATES, OPTIONS, VARIABLES, DEFINITIONS. And note that a section header is always placed by itself on a line and has a colon appended to it, as in TEMPLATES:.
- 3. The TEMPLATES section of the file contains template definitions for formatting references. The SPECIAL-TEMPLATES section contains definitions for creating variables within each database entry. The OPTIONS section contains definitions for various keywords that can be used to modify program behavior. The VARIABLES section contains user definitions for new variables to be made available. The difference between these definitions and those in SPECIAL-TEMPLATES is that the ones provided in the VARIABLES section are in-line Python code, whereas the former use templates. Finally, the DEFINITIONS section of the file contains Python-executable code that can then make functionality available in the form of template variables. (An example is provided in the *Python API* section below.
- 4. All variable definitions within the TEMPLATES, SPECIAL-TEMPLATES, and OPTIONS sections use the variable name followed by whitespace, an equals sign, more whitespace, and then the definition itself. Thus the [whitespace+=+whitespace] expression is the delimiter between variable and definition, and is required syntax.
- 5. An ellipsis . . . at the end of a line indicates a line continuation. All whitespace following the ellipsis, and all whitespace preceding text on the next line, is removed from the resulting connected text.
- 6. An ellipsis in the middle of a line indicates an "implicit loop". For details, see the *Examples for namelist formatting* section below.
- 7. Inside the TEMPLATES section, all of the variable definitions are intended to map to a database entrytype name. For example:

```
article = <au>, ``<title>,'' <journal>, <volume>, <startpage>--<endpage> (<year>).
```

Here the article entrytype will be typeset so that the list of authors (<au>) is followed by the article title in double quotes, the journal name in standard font (i.e. not italics), the volume number, the page range, and the year.

8. A variable is indicated by angle brackets, as <var> and represents the contents of a field found within the bibliography database. Thus, when typesetting the bibliography, Bibulous will replace the variable <authorlist> with the string stored in the authorlist field of the current entry being formatted. An example list of typical variables one might use is:

```
<au>, <booktitle>, <chapter>, <edition>, <ed>, <eid>, <endpage>,
<institution>, <journal>, <nationality>, <note>, <number>,
```

```
<organization>, <publisher>, <school>, <series>, <startpage>, <title>,
<version>, <volume>, <year>.
```

This list is actually freely extensible. A user can add any additional variables needed, so that if a video field is used in a .bib database file, then this can be used within a formatted reference simply by placing <video> into the template wherever the information needs to be inserted.

- 9. Any variable placed within square brackets [] indicate that it is an optional variable it is not required that the database have that entry. While required entries that are not defined in the BibTeX database file (.bib file) are replaced with '???', and undefined optional variables are simply skipped. If a | is present within the square brackets, it indicates an "elseif" clause. That is, if the template is [<var1>|<var2> and <var3>], then the code will look for var1 as a field within the current database entry being formatted. If it does not find the entry, then it will try the next block, where it finds the two variables var2 and var3. If both are defined, then the original template [<var1>|<var2> and <var3>] is replaced with <var2> and <var3> (i.e. removing the square brackets) and proceeds to replace the two variables with their corresponding fields. If either one of var2 or var3 is undefined in the entry, then the entire optional (bracketed) portion of the template is skipped.
- 10. If the | symbol is used to create an empty last cell, as in [<var1>|<var2>|], this indicates that while the individual cells within the optional block are themselves optional, it is required to have at least _one_ among the cells to be defined. Thus, [<note>|] has the same meaning as simply <note> does.
- 11. Nesting of [] brackets is allowed, but the syntax becomes computationally expensive to parse, so that these structures should be used sparingly.
- 12. Users that need to use [,], #, <, >, or | symbols as formatting elements within the reference list can implement them using some custom LaTeX-markup commands: {\makeopenbracket}, {\makeclosebracket}, {\makeclosebracket}, or {\makeellipsis}. Note that the curly brackets used in each case are required.
- 13. In the TEMPLATES section of the file, if an entrytype format definition contains only another entrytype name on the right hand side of the =, i.e.:

```
inbook = incollection
```

then this means that the existing incollection template should be copied for use with inbook entrytypes.

- 14. Note that several fields are defined by default which are *not* directly from the bibliography database. These are au, authorlist, citekey, citenum, ed, editorlist, endpage, and startpage. These fields are derived from the original database file, but have been reformatted. See the *Default Fields* section below.
- 15. Although the entrytype template definitions listed below are in alphabetical order, that can be put in any desired order within the file. The exception to this rule is that if a definition consists of, for example:

```
inbook = incollection
```

then the incollection template must already be defined. Also note that two entrytype names are special and so cannot be used on the left hand side of the equals sign here: comment and preamble.

- 16. A user wanting a localized form of quotation should use \enquote { <title>} rather than `` <title>'', and add \usepackage { csquotes } to the preamble of the LaTeX document.
- 17. In the OPTIONS section of the file are the formatting options. None of these definitions are required. The complete list is given in the *Options keywords* section below, together with explanations of each.
- 18. The SPECIAL-TEMPLATES section is where users can define their own fields that get generated for every database entry evaluated. For example, the variable definition:

```
group = [<organization>|<institution>|<corporation>|]
```

in the SPECIAL-TEMPLATES section will create a group field that can be used as the variable <group> within the regular TEMPLATES section of the file. This effectively allows users to create a shortcut. Where they could write out [<organization>|<institution>|<corporation>|] inside each template that needs this structure, the special template definition allows them to replace each instance with a simple <group>.

Another example usage would be the following:

```
author = [<author-en>|<author-jp>|]
```

where the author field is actually redefined to include not only the *existing* author field, but also the fields author—en or author—jp. That is, if the author field is missing in the database entry (the field matching the thing on the left hand side), the code next searches for the author—en field. If it finds it, then it will create an author field that contains a copy of the author—en's field's contents. If the author—en is also missing, the code next searches for author—jp and uses that fields contents to create the missing author field. This is a convenient way of grouping different variable names in order to simplify templates.

19. The order in which any definitions are placed within the special templates is important. For example, if a user has au = <authorlist.format_authorlist() > and then below that defines authorlist = <author.to_namelist() >, then the code will issue an error stating that authorlist is not defined when attempting to create the au variable. Since the definition for au assumes the presence of the authorlist variable, the latter definition must be placed above it.

3.2 Default Fields

A complete list of the existing default fields is::

```
au
authorlist
citealnum
citealpha
citekey
citenum
ed
editorlist
sortkey
sortnum
```

Each of these default fields are defined as "special templates". If a user defines a special template with the same name as one of the above, then the default is overwritten with the user's version. The definitions of the default special templates are::

```
authorlist = <author.to_namelist()>
editorlist = <editor.to_namelist()>
citealnum = ?
citelabel = <citenum>
sortkey = <citenum>
au = <authorlist.format_authorlist()>
ed = <editorlist.format_editorlist()>
```

Note that the ordering of definitions is important (that is, if citelabel uses authorlist in its definition, then the latter has to appear before it in the list of special templates, so that it is defined by the time that citelabel needs to be defined). The following summarizes what these definitions are used for.

Also, in addition to these default fields, there are also four special variables defined, citealpha, citealnum, citenum, and sortnum. These can be accessed as <citealpha>, <citealnum>, <citenum>, and <sortnum> within any template. The first, <citealpha>, is designed to reproduce the citation label style used

3.2. Default Fields

by BibTeX's alpha style. For example, GKP94 is the alpha—style citation label for the book *Concrete Mathematics*, 2nd Edition, by Ronald L. Graham, Donald E. Knuth, and Oren Patashnik (1994). The second variable <citealnum> is a variant of this that creates a citation label from the first letter of the author's last name, followed by an integer indicating its place in the sorted bibliography. Thus the book by Graham *et al.* would have the label "G2" if preceded by, say, *Introduction to Fourier Optics* by Joseph W. Goodman (1968), since the latter, being published earlier, would be placed earlier in the sorted list. <citenum> is simply the numerical order in which the reference was cited in the text (starting at 1 with the first citation). <sortnum> is the numerical order in which the citation appears in the reference list. If the reference list is sorted alphabetically, for example, then sortnum will generally be different from citenum.

au is the string representing the formatted list of author names. In the default definition shown above, the name list is a standard form, and so simply uses the .format_authorlist() operator. Generally, this operator creates name lists that have the form "firstauthor" for only one author, "firstauthor and secondauthor" if only two authors, "firstauthor, secondauthor, ..., and lastauthor" if more than two authors but less than the maximum, and "firstauthor, secondauthor, ..., minauthor, et al." if more than the maximum allowed number of authors. Which author in the list is "minauthor" is defined using the minauthors option keyword. The maximum number of allowed authors is set by the maxauthors option keyword.

authorlist creates a list of dictionaries (one dictionary for each author name found within the database entry's author field). Each name dictionary has keys "first", "middle", "prefix", "last", and "suffix", where each of these keys is optional except for "last". Thus, a user can access the first and last name of the first author in the database entry using <authorlist.o.first> <authorlist.o.last>. To access the middle name(s) of the second author, use <authorlist.l.middle>.

citealnum is a field generated by taking the first letter of the first author's last name, and appending to it the citation sort order number within all entries sharing that prefix. Thus "Bugs Bunny" would have a citealnum of "B" followed by its order number. If "Yogi Bear" were another first-author name in the list, then "Yogi Bear" would receive citealnum = B1 while "Bugs Bunny" would receive citealnum = B2.

citealpha is a field generated according to BibTeX's classic "alpha" style. That is, it takes the first three letters of the author's last name if there is only one author. If multiple authors, it takes the initials of the author last names, for up to the first three authors. If no author is present, then it looks for the first three letters of the organization field. Finally, the three-letter alphabetical label is followed by the last two digits of the publication year. Thus, any user wishing to use BibTeX's classic alpha style for citation labels can use citelabel = <citealpha> in the SPECIAL-TEMPLATES section of the file.

citelabel is the thing that appears at the front of the formatted reference, and is identical to the citation label used in the manuscript to point to the item in the reference list. In technical journal articles, this is typically just a number, as in the default definition <citenum>. The number used here for the label indicates the order in which the entry was cited.

ed follows the same basic structure as au, but uses the maxeditors and mineditors keywords.

editorlist behaves exactly as authorlist but derives its list of names from the database entry's editor field rather than author field.

sortkey is the string used to sort the entry within the reference list. For technical journal articles, what is generally wanted is just the citation order, as indicated by the <citenum> variable.

3.2.1 Operators

One can use the "dot" operator inside a variable name, as in <authorname.0.last.initial()>'' to perform any one of five functions: a explicit numerical index (the ``O shown here, listed as the .# operator below), an implicit numerical index (using .n or .N, for which see section *Examples for namelist formatting* below for details), a range index (listed as the ##:## operator below), a dictionary lookup (the last used here), or the application of an operator (in this case, the .initial() operator which is used to reduce a name to its initial). A numerical index must apply to a list-type of variable, and a key index must apply to a dict-type of variable (i.e. a dictionary). Note that, due to limitations of the parser, it is not allowed for <.>

angle-bracketed variables to be placed within other angle-bracketed variables. Thus, function arguments take variable names with their angle brackets removed.

The complete list of operators available is::

```
. #
.##:##
.compress()
.format_authorlist()
.format_editorlist()
.frenchinitial()
.if_len_equals(var1, number, var2, var3)
      .if_len_less_than(var1, number, var2, var3)
      .if len more than (var1, number, var2, var3)
.if_str_equal(test_str,then_var,else_var)
.if_singular(var1, var2, var3)
.initial()
    .len()
.lower()
.monthabbrev()
.monthname()
. N
.null()
.ordinal()
.purify()
.replace(old, new)
.sentence_case()
.tie()
.to_namelist()
.uniquify(arg)
.upper()
.zfill(num)
```

The function of each operator is summarized below.

.# An explicit numerical index, i.e., select the #th element of the operand.

.##:## A range index, i.e., select the ##th through ##th elements of the operand. For example, for a bibliography entry whose database file contains title = {Impossibility}, a template variable of the form <title.0:2> will return Imp, and <title.3:5> will return oss. The first character of the operand thus has an index 0. Indexing from the end of the operand can be done using negative number indices. For example, the last character of the operand can be indexed by -1, and the third to last by -3, so that in the example above, <title.-3:-1> will return ity.

.compress() removes any whitespace found within the string. This is useful for generating namelists where the format requires "tight" spacing. An example would be "RMA Azzam", where the three initials are grouped together without spacing. And example template for generating this type of name would be:

```
<authorlist.0.first><authorlist.0.middle.initial().compress()> <authorlist.0.last>
```

Without the .compress() operator, the name would come out as "RM A Azzam", since the two middle name initials "M" and "A" are spaced apart from one another by default.

.format_authorlist() operates on a list of dictionaries type of variable (a namelist), and uses the keyword-based default formatting scheme to create a formatted string of names. The complete list keywords that it looks for is: etal_message, maxauthors, minauthors, namelist_format, period_after_initial, terse_inits, use_firstname_initials, use_name_ties. The default formatter, while fast, is not very flexible, so that users looking for more customizability will want to make use of Bibulous' implicit-index and implicit-loop based definitions. See the *Example definitions for namelist formatting* section below.

3.2. Default Fields 15

.format_editorlist() operates on a list of dictionaries type of variable (a namelist), and uses the keyword-based default formatting scheme to create a formatted string of names. The complete list keywords that looks for is: etal_message, maxeditors, mineditors, namelist_format, period_after_initial, terse_inits, use_firstname_initials, use_name_ties. (The difference with the .format_authorlist() operator is that it uses maxeditors and mineditors rather than maxauthors and minauthors) The default formatter, while fast, is not very flexible, so that users looking for more customizability will want to make use of Bibulous' implicit-index and implicit-loop based definitions. See the Example definitions for namelist formatting section below.

.frenchinitial() is an alternative form of the .initial() operator that has slightly different behavior. If a name begins with one of the digraphs "Ch", "Gn", "Ll", "Ph", "Ss", or "Th", then the initial will truncate the name after the digraph instead of after the first letter.

.if_str_equal(test_str,then_var,else_var) inserts then_var if the string to be operated on is equal to test_str,
else it inserts the string else_var.

.if_singular(var,res1,res2) inserts res1 if var has only one element, but res2 if var has more than one element. Here var is assumed to be a list-type of variable. An example usage is the following: [<ed. if_singular(editorlist, option.edmsg1), option.edmsg2)>,]. This form appends the contents of the edmsg1 option variable to the end of the ed variable if the editorlist variable contains only one element (is singular). If editorlist has more than one element, then the contents of the edmsg2 variable is appended instead. (This could potentially be a null string.)

.if_num_equals(var,number,res1,res2) appends res1 if var equals number, but otherwise appends var2.

.initial() will truncate the string to its first letter. Note that if a name begins with a LaTeX markup character, such as $\{\'E\}$, then the operator will convert the input string to its best attempt at a Unicode-equivalent (without character markup) prior to performing the truncation. Thus, applying the .initial() operator to the name $\{\vE\}\}$ ukauskas will produce the initialized form "Ž".

.len() returns the number of elements in a list variable (to the left of the operator dot).

.lower() reduces all letters in its argument to lower case. If any LaTeX markup letters exist in the argument, then they will be converted to Unicode equivalents first before applying the operator. Thus, if the field au contains {\AA} then the operator will first convert this to the letter Å and then reduce it to the lower case å.

.monthabbrev() assumes that the input field is a number from 1 to 12, and converts the numerical input into the abbreviated month according to the user's current locale. If the system cannot determine the user's locale, the operator will default to using the American English locale, which replaces the numerical field operated on with one of "Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", or "Dec" according to the field's value. Thus, if the bibliography database entry has a field month = 11, and the template has the form <month.monthabbrev()>, then the template will be replaced with "Nov" for the default locale. For users with locale "Japan", this same operator will return "11".

.monthname() behaves much like .monthabbrev() but rather than using an abbreviated form for the month's name, it uses the full form. Thus if the bibliography database entry has a field month = 3, and the template has the form <month.monthname()>, then the template variable will be replaced with "March" for the default locale. For users with locale "Norway", this same operator will return "Mars".

.n See the Examples for namelist formatting below

.N [Not currently supported]

.null() returns an undefined variable, forcing the string to be evaluated as "undefined".

.ordinal() creates an "ordinal" from a numerical field. Thus, if the field operated on is "1", "2", "3", or "4", then the operator will replace the template with "1st", "2nd", "3rd" or "4th". Any number above 4 simply has "th" appended to the end of it. Currently Bibulous does not support non-English locales for this function. (Anyone having suggestions of how this may be implemented without too much fuss should contact us!)

.purify() attempts to convert its argument into a string without LaTeX-markup for foreign characters. Thus, if the entry contains title = $\{\{AA\}\$ then a template variable of the form <title.purify() > will produce the result Aland. This can be useful when having to use $substr_replace()$ and other functions where the markup may cause matching problems.

.replace(old,new) will replace the substring old with new wherever it finds old within the string it is applied to. For example, if a user wants to make the name "J. W. Tukey" bold everywhere it appears in a reference, then .replace(J. W. Tukey, \textbf{J. W. Tukey}) will work. Note that whitespace is preserved here. Thus, .replace(J. W. Tukey, \textbf{J. W. Tukey}) will add a space in front of \textbf{J. W. Tukey}. Also, Bibulous will not allow the use of <, >, |, or) characters in the two arguments of the operator.

.sentence_case() reduces the lower case any characters in the field, except for the initial letter and any letters protected
within a pair of curly braces. For example, if the database entry has title = {Understanding Bohmian
mechanics} and the template has the form <title.sentence_case()>, then the template variable will
be replaced with "Understanding bohmian mechanics". However, if the entry has title = {Understanding
{B}ohmian mechanics}, the result will be "Understanding {B}ohmian mechanics".

.tie() replaces any spaces with an unbreakable space. Thus, "R. M. A." becomes "R.~M.~A.". An example use of this operator would be the following template::

```
authorname = [<authorlist.n.first.initial()>.~][<authorlist.n.middle.initial() .tie()>.

→ ]...

[<authorlist.n.prefix>~]<authorlist.n.last>[, <authorlist.n.suffix>]
```

.to_namelist() parses the field (assumed to be a BibTeX-format "and"-delimited list of names) into a Bibulous-format namelist (i.e. a list of dictionaries).

.uniquify(arg) appends a letter character (if arg="arg" or a number (if arg="1") to the end of the field to make it unique relative to the same field in every other cited entry. For example, if two authors have the name "Smith" and published a paper in the year 2000, then the template citelabel = <authorlist.0.last><year> will produce the same citation label for both entries. Thus, we can rename this as citetemp = <authorlist.0.last><year> and then define citelabel = <citetemp.uniquify(a)> to generate unique labels. For this example, the first citation will have the label "Smith2000" and the second "Smith2000a". If a third citation shares the same name and year, then it will be given the unique label "Smith2000b", and so on.

.upper() raises all letters in its argument to upper case. If any LaTeX markup letters exist in the argument, then they will be converted to Unicode equivalents first before applying the operator. Thus, if the field au contains {\aa} then the operator will first convert this to the letter å and then raise it to the upper case Å.

.zfill(num) appends zeros to the front of its argument, where num indicates the desired final length of the string. For example, if the field vol contains the number 11, then calling <vol.zfill(3) > produces the result 011.

3.3 Options keywords

A complete list of existing options keywords, together with their default definitions, is::

```
allow_scripts = False
autocomplete_doi = True
backrefs = False
backrefstyle = none
bibitemsep = None
case_sensitive_field_names = False
edmsg1 = , ed.
edmsg2 = , eds
etal_message = , \\textit{et al.}
maxauthors = 9
maxeditors = 5
```

(continues on next page)

(continued from previous page)

```
minauthors = 9
mineditors = 5
name_separator = and
namelist_format = first_name_first
period_after_initial = True
procspie_as_journal = False
sort_case = True
sort_order = forward
terse_inits = False
undefstr = ???
use_abbrevs = True
use_citeextract = True
use_firstname_initials = True
use_name_ties = False
```

Each of the keywords is summarized below.

allow_scripts [default value: False] tells Bibulous whether to allow the evaluation of Python code in the VARIABLES and DEFINITIONS sections of .bst files. It is important for users to realize that evaluating external code in this way is a security risk, and so they should not set allow_scripts = True when inserting code that they do not trust. However, as an additional security precaution, Bibulous prevents most security-sensitive operations from being used within its Python API.

autocomplete_doi [default value: True] tells Bibulous whether to add http://dx.doi.org/ to the front of the doi field if the front is missing. This allows the <doi> variable to be used as a complete URL, even when the prefix is missing in the database field.

backrefs [default value: False] THIS KEYWORD IS NOT YET IMPLEMENTED

backrefstyle [default value: none] THIS KEYWORD IS NOT YET IMPLEMENTED

bibitemsep [default value: None] provides users a means to change the amount of vertical separation that LaTeX sets between entries in the reference list. For example, users wanting a more compact list can define bibitemsep = 0pt.

case_sensitive_field_names [default value: False] tells Bibulous whether to consider, for example, a field named "Author" as being distinct from "author".

edmsg1 [default value: , ed.] provides a string to use after a list of editor names, for the case when only one editor is present.

edmsg2 [default value: , eds] provides a string to use after a list of editor names, for the case when multiple editors are present.

etal_message [default value: , \textit{et al.}] provides a string to use after a truncated namelist (for example, when the number of authors exceeds the value given by the maxauthors keyword).

maxauthors [default value: 9] provides the maximum number of allowed names in the formatted list of authors. If the number of names is more than this, then the list of names is truncated to minauthors and the etal_message is appended to the result. (This keyword is only used within the .format_namelist() operator.)

maxeditors [default value: 5] provides the maximum number of allowed names in the formatted list of editors. If the number of names is more than this, then the list of names is truncated to mineditors and the etal_message is appended to the result. (This keyword is only used within the .format_namelist() operator.)

minauthors [default value: 9] provides the minimum number of author names to use when truncating an overlength author name list. (This keyword is only used within the .format_namelist() operator.)

mineditors [default value: 5] provides the minimum number of editor names to use when truncating an overlength author name list. (This keyword is only used within the .format_namelist() operator.)

name_separator [default value: and] informs Bibulous how to separate the individual names in a BibTeX-format field of names. For example, with a BibTeX-format field of Bugs Bunny and Porky Pig, using the separator and allows Bibulous to determine that there are two separate people, Bugs Bunny and Porky Pig. If the option keyword has an empty field (*i.e.* it is written as name_separator =, with nothing on the right hand side of the equals) then it is assumed that the intended separator is a space character. This is useful for alphabets (such as Chinese and Japanese) that often work without spaces.

namelist_format [default value: first_name_first, allowed values: {first_name_first, last_name_first}] defines how the
formatted list of names should appear. If namelist_format = first_name_first then the individual names
will appear in the order "firstname middle prefix last, suffix". If namelist_format = last_name_first then
the individual names will appear in the order "prefix last, firstname middle, suffix". (This keyword is only used within
the .format_namelist() operator.)

period_after_initial [default value: True] tells the .format_namelist() operator whether to place a period
after each initial of an individual's name. Thus, if period_after_initial = True, a name will appear
as "R. M. A. Azzam", but if False will appear as "R M A Azzam". (This keyword is only used within the .
format_namelist() operator.)

procspie_as_journal [default value: False] The "Proceedings of SPIE" are treated as special by the journals of the Optical Society of America. That is, they format these proceedings (and only these) in the same way that they do journal articles. Thus, a special keyword is required to allow this behavior.

sort case [default value: True] informs Bibulous whether or not to use case-sensitive sorting of reference keys.

sort_order [default value: forward] whether to sort in increasing order (forward) or decreasing order (reverse).

terse_inits [default value: False] tells the .format_namelist() operator whether to compress together the initials of an individual's name. Thus, if terse_inits = True, a name will appear as "RMA Azzam", but if False will appear as "R. M. A. Azzam". (This keyword is only used within the .format namelist() operator.)

undefstr [default value: ???] informs Bibulous what kind of warning message to print when a required field is missing in the database entry.

use_abbrevs [default value: True] tells Bibulous whether or not to use the abbreviations defined in the bibliography database. (Used for debugging.)

use_citeextract [default value: True] tells Bibulous whether to perform "citation extraction", which creates a small database of only the cited items from among the complete database provided in the .aux file.

use_firstname_initials [default value: True] Whether or not to initialize the first names of authors in the formatted
authors list. (This keyword is only used within the .format_namelist() operator.)

use_name_ties [default value: False] Whether or not to replace spaces with unbreakable spaces (i.e. "R. M. A. Azzam" or "R.~M.~A. Azzam") inside names in the name list. (This keyword is only used within the . format namelist() operator.)

3.4 Implicit loops and examples for namelist formatting

The following code provides an example usage of implicit indexing within an implicit loop structure::

(continues on next page)

(continued from previous page)

```
[<editorlist.n.prefix> ] <editorlist.n.last>[, <editorlist.n.suffix>]
ed = <editorname.0>, ..., { and } <editorname.2>
```

Here the authorlist and editorlist definitions create namelist variables from the author and editor fields in the entry (if they exist). Next, the implicitly-indexed authorname.n cannot operate except within an implicit loop, and so we should describe that first. It is easier to describe the functionality of the ed template than the au one, as it has a smaller number of allowed names. The ed template has the definition:

```
<editorname.0>, ...,{ and }<editorname.2>
```

which simplifies to <editorname.0> when there is only one editor in the database entry, and:

```
<editorname.0> and <editorname.1>
```

when there are only two. Here the separator "and "comes from the { and } placed at the right hand side of the implicit loop. For three editors, the implicit loop expands the template to:

```
<editorname.0>, <editorname.1>, and <editorname.2>
```

where this time the comma alone is used as the first delimiter, as it is outside the enclosed braces. For the final element, both the comma and the { and } at the right hand side of the implicit loop are used as the final delimiter. Since the template does not specify the format for more than three editor names, the code builds an *et al.* construction when there more than this number of names, so that the result becomes:

```
<editorname.0>, <editorname.1>, <editorname.2>, \textit{et al.}
```

where the form of the string \textit{, et. al} is specified by the etal_message keyword option.

Thus, the implicit loop has filled out a unique template based on the number of editors it finds within the database entry. The next step is to use the implicitly-indexed editorname to complete building out the template. The latter template is defined as:

so that a template variable of the form "<editorname.0>" is replaced with:

```
[<editorlist.0.first.initial()>. ][<editorlist.0.middle.initial()>. ]...
[<editorlist.0.prefix> ]<editorlist.0.last>[, <editorlist.0.suffix>]
```

That is, the implicit index .n is everywhere replaced with the explicit index 0. For the case of a database entry containing two editor names, the final template will thus have the form:

```
[<editorlist.0.first.initial()>. ][<editorlist.0.middle.initial()>. ]...
[<editorlist.0.prefix> ]<editorlist.0.last>[, <editorlist.0.suffix>] and ...
[<editorlist.1.first.initial()>. ][<editorlist.1.middle.initial()>. ]...
[<editorlist.1.prefix> ]<editorlist.1.last>[, <editorlist.1.suffix>]
```

With this template now complete, the code begins to evaluate the entry and replace the individual variables with their corresponding database fields.

3.5 Python API

Bibulous also provides to users an extensible Python interface allowing users to directly manipulate Bibulous' internal data structures. These use the VARIABLES and DEFINITIONS sections of the file, as shown below. For the

`VARIABLES section, a variable name is defined (the first example below defines the variable year_bce, while the second example below defines pagerange). On the right hand side of the definition, however, is a Python function call. This is different from the other sections of the BST file, which use template syntax. Any variable defined in this way within the VARIABLES section can then be accessed as a template variable (i.e. <year_bce>) within the TEMPLATES section of the file. Two example uses are shown below.

To allow Bibulous to read the VARIABLES and DEFINITIONS sections of the file, users must set the option keyword allow scripts to True.

First example: a custom yearstyle. For a bibliography containing works from authors dating from before year 0, a common approach is to append "BC" to the year number, and for positive-numbered years, appending "AD". More recently, the convention has been to append "BCE" and "CE" rather than "BC" and "AD". The example defines an option keyword yearstyle that allows users to switch between one style (BC/AD) and the other (BCE/CE). This keyword is accessed by placing options as an argument to the format_yearstyle() function defining the variable year_bce. Inside the function, it can then check the options dictionary for the yearstyle keyword and determine which convention to use.

The format_yearstyle() function itself is straightforward. It first checks whether the entry has a year field. If not, then it returns None, indicating that the function's result is undefined. If it finds a year field, then it checks to see whether it corresponds to an integer. If not, then it returns the field as-is. (Perhaps a user defines his year fields as 45 BCE with the BCE already written out inside the field?) If it finds an integer value, then it determines which style to use (BC/AD or BCE/CE). If the year number is negative then it appends "BC" or "BCE to the end. If the year number is positive then it appends "AD" or "CE to the end, depending on the convention chosen.

Example:

```
OPTIONS:
allow_scripts = True
yearstyle = BCE/CE
VARIABLES:
year_bce = format_yearstyle(entry, options)
DEFINITIONS:
## NOTE! Only Unix-style line endings are allowed here.
def format_yearstyle(entry, options):
   Append "BC or "AD" to "year", depending on whether the year is positive or,
⇔negative.
   If the option "yearstyle" is set to "BCE/CE", then use "BCE" and "CE" instead of
→ "BC"
    and "AD".
    1.1.1
    if ('year' not in entry):
        return(options['undefstr'])
    ## First check that the year string is an integer. If not an integer, then just_
→ return
    ## the field itself.
    if not str_is_integer(entry['year']):
        return (entry['year'])
   yearnum = int(entry['year'])
    if (yearnum < 0):</pre>
        if (options['yearstyle'] == 'BCE/CE'):
            suffix = 'BCE'
```

(continues on next page)

3.5. Python API

(continued from previous page)

```
else:
    suffix = 'BC'
## The "[1:]" here removes the minus sign.
    result = str(yearnum)[1:] + ' ' + suffix

elif (yearnum == 0):
    result = str(yearnum)

else:
    if (options['yearstyle'] == 'BCE/CE'):
        suffix = 'CE'
    else:
        suffix = 'AD'
    result = str(yearnum) + ' ' + suffix
return(result)
```

Second example: a custom pagestyle. For a bibliography containing works from magazines, it is not uncommon to find articles with large gaps in page numbers. Here is an example bibliography database entry:

```
@article{stewart,
title = {Interview with Walter Stewart},
author = {Doug Stewart},
journal = {Omni},
year = {1989},
volume = {11},
number = {5},
pages = {64--66, 87--92, 94}
}
```

where we can see that the article was broken into three sections in order to fit the editors' formatting requirements. Many bibliography styles require a starting and ending page, but these are misleading when the article is broken across pages in this way. Thus, a user may want to have the option that if a comma is found within the pages field of an entry then it should be displayed as-is. If no comma is found, then it simply returns the standard startpage—endpage pair.

To make this work, first the option allow_scripts must be set to true. Next, a new pagerange variable is defined, so that it can be accessed in the TEMPLATES section of the file as <pagerange>. The variable is defined as the return value of the function format_pagerange() given in the DEFINITIONS section. The defined function first checks to see if there is a pages field defined in the entry. If not, then it returns None, so that the pagerange variable will also be undefined. If it finds the pages field, it looks to see if there is a comma present. If so, it returns the field as-is. If not, it looks for the endpage variable (generated by default by Bibulous from the pages field). If present, then the function returns a startpage—endpage pair. If endpage is not present, then it returns only the startpage variable.

Example:

(continues on next page)

(continued from previous page)

```
return the standard startpage--endpage range.
'''

if not ('pages' in entry):
    return(None)

elif (',' in entry['pages']):
    return(entry['pages'])

elif ('endpage' in entry):
    return(entry['startpage']--entry['endpage'])

else:
    return(entry['startpage'])
```

3.5. Python API

INSTRUCTIONS ON HOW TO REPORT A BUG TO THE BIBULOUS DEVELOPMENT TEAM

4.1 Where to report a bug

Send an email to the users_mailing_list. Once it's confirmed as a bug, someone, possibly you, can enter it into the issue tracker. (Or if you're pretty sure about the bug, go ahead and post directly to the developers mailing list, developers_mailing_list. But if you're not sure, it's better to post to [users mailing list] first; someone there can tell you whether the behavior you encountered is expected or not.)

4.2 How to report a bug

First, make sure it's a bug. If Bibulous does not behave the way you expect, look in the documentation and mailing list archives for evidence that it should behave the way you expect. If the documentation and archives do not contain enough information to tell you whether the behavior is a bug or is expected behavior, go ahead and ask on the users mailing list first users_mailing_list. Also check that you are running the most recent version of Bibulous. It may be that the bug has already been fixed.

Once you've established that it's a bug, the most important thing you can do is come up with a simple description and reproduction recipe. For example, if the bug, as you initially found it, involves five files over ten commits, try to make it happen with just one file and one commit. The simpler the reproduction recipe, the more likely a developer is to successfully reproduce the bug and fix it.

When you write up the reproduction recipe, don't just write a prose description of what you did to make the bug happen. Instead, give a copy of the exact series of commands you ran, and their output. Use cut-and-paste to do this. If there are files involved, be sure to include the names of the files, and even their content if you think it might be relevant. The very best thing is to package your reproduction recipe as a script, that helps a lot.

In addition to the reproduction recipe, we'll also need a complete description of the environment in which you reproduced the bug. That means:

- Your operating system
- The Python version you are running under.
- The release and/or revision of Bibulous.
- Anything else that could possibly be relevant. Err on the side of too much information, rather than too little.

Once you have all this, you're ready to write the report. Start out with a clear description of what the bug is. That is, say how you expected Bibulous to behave, and contrast that with how it actually behaved. While the bug may seem obvious to you, it may not be so obvious to someone else, so it's best to avoid a guessing game. Follow that with the environment description, and the reproduction recipe. If you also want to include speculation as to the cause, and even suggest how the code may be modified to fix the bug, that's great.

Post all of this information to the developers mailing list, developers_mailing_list, or if you have already been there and been asked to file an issue, then go to the Issue Tracker and follow the instructions there.

Thanks! It's a lot of work to file an effective bug report, but a good report can save hours of a developer's time, and make the bug much more likely to get fixed.

CHAPTER

FIVE

DEVELOPER GUIDE

5.1 Guidelines and notes for Python coding style

- 1. Note that one can mix 8-bit Python strings (ASCII text) with UTF-8 encoded text as long as the 8-bit string contains only ASCII characters.
- 2. Keep in mind when running into Unicode errors: reading a line of text from a file produces a line of bytes and not characters. To decode the bytes into a string of characters, you need to know the encoding.
- 3. There are a couple of minor points where the Bibulous coding standards deviates from Python's PEP8:
 - (a) A line width of 120 is the standard (not 80).
 - (b) In general, statements that evaluate to a boolean are placed within parentheses (i.e. if (a < b): rather than if a < b:).
- 4. Many developers prefer to spread out code among a large number of small files, but Bibulous is currently organized in a single large file. This is partly because there is no large block of code that fits by itself so that a separate file makes sense. (Parsing of .bib files, for example, only requires a couple hundred lines.

5.2 Overall project strategy and code structure

The basic function of BibTeX is to accept an .aux file as input and to produce a .bbl file as output. The aux file contains all of the citation information as well as the filenames for the bibliography database file (.bib) and the style file (.bst).

The basic program flow is as follows:

- 1. Read the .aux file and get the names of the bibliography databases (.bib files), the style templates (.bst files) to use, together with the entire set of citations.
- 2. Read in the Bibulous style template file as a dictionary (bstdict).
- 3. If the use_citeextract keyword is set to True, and if an "extracted" database file exists, then compare the citations in the extracted database against those in the .aux file. If there are any differences, then re-extract the database. Otherwise, use the extracted database rather than the full one specified in the .aux file.
- 4. Read in all of the bibliography database files into one long dictionary (bibdata), replacing any abbreviations with their full form. In an "extracted" database, all entries are parsed, whereas in any other type of database file, only those entries whose keywords are found in the citation list are actually parsed. All other entries have their data saved as unparsed strings. Cross-referenced data is *not* yet inserted at this point. That is delayed until the time of writing the BBL file in order to speed up parsing. It is only then that the cross-referenced entries have their data parsed into dictionary form.
- 5. Now that all the information is collected, we can generate the .bbl file. Create the list of sortkeys, then go through each corresponding citation key in turn, and find the corresponding entry key in bibdata. If there is

crossref data, then fill in missing values here. Also create the "special variables" here. Finally, from the entry type, select a template from bstdict and begin inserting the variables one-by-one into the template.

Because the .bib file is highly structured, it is straightforward to write a parser by hand in Python: the parse_bibfile() method converts the .bib file contents into a Python dictionary (the Bibdata class' bibdata). The .aux file is even easier to parse, and the parse_auxfile() method converts the citation information into the Bibdata class' citedict dictionary.

The Bibdata class thus holds all relevant information needed to operate on a bibliography and generate the output LaTeX-formatted .bbl file.

5.3 Parsing BIB files

5.3.1 parse_bibfile()

The strategy for parse_bibfile() is to find each individual bibliography entry, determine its entry type, and save all of the text between the entry's opening and closing braces as one long string, to be passed to parse_bibentry() for further parsing. To gather the entry data string, we first look for a line that starts with @. On that line, we look for a string after the @ followed by {, where the string gives the entry type. After we know the entry type, we look for the corresponding closing brace. If we don't find it on the same line, then we read in the next line, and so forth, concatenating all of the lines into one long "entry string" until we encounter the corresponding closing brace. Once we have this extended "entry string" we feed it to parse_bibentry() to generate the bibliography data. Once we have come to the end of a given entry, we continue reading down the file looking for the next '@' and so on.

Although this approach effectively means that we have to pass twice through the same data, dealing with brace-matching can otherwise become a mess for the BibTeX format, since it allows nested delimiters, is not directly compatible with regular expressions.

5.3.2 parse_bibentry()

Because parse_bibfile() has already split the data by individual entry, parse_bibentry() only needs to worry about parsing a single entry, and there are five possible formats for the entry string passed to the function:

- 1. If the entrytype is a comment, then skip everything, adding nothing to the database dictionary.
- 2. If the entrytype is a preamble, then treat the entire entry contents as a single fieldvalue. Append the string onto the preamble value in the bibdata dictionary.
- 3. If the entrytype is a acronym, then get the entrykey and copy it into the name field. The remainder of the string is a single field value (the full form of the acronym); copy that into the description field.
- 4. If the entrytype is a string (i.e. an abbreviation), then there is no entrykey. Get the fieldname (abbreviation key), and the remainder of the string is a single field value (the full form of the abbreviated string). Add this key-value pair to the abbrevs dictionary.
- 5. If the entry is any other type, then get the entrykey, and the remainder of the string is a series of field-value pairs.

Once it determines which of these four options to use, parse_bibentry() extracts the entry key (if present), it locates each individual field and separates out the string corresponding to the key-value pair for each field. It does not actually *parse* the individual fields. For that, it loops over each field with a call to parse_bibfield() to extract the field key-value pairs.

5.3.3 parse bibfield()

parse_bibfield() is the workhorse function of the BIB parsing. And because of BibTeX's method for allowing concatenation, use of abbreviation keys, and use of two different types of delimiters ("..." or {...}), this function is a little messy. However, for the format of a given field, there are four parsing possibilities:

- 1. If the field begins with a double quote " then scan until you find the next unnested ". Add that to the result string. If the ending " is followed by a comma, then the field is done; return the result string. If the ending is followed by a # then expect another field string. Scan for it and append it to the current result string.
- 2. If the field begins with { then scan until you resolve the brace level. This should be followed by a comma, since no concatenation is allowed for brace-delimited fields. Otherwise issue a syntax error warning.
- 3. If the field begins with a # (concatenation operator) then skip whitespace to the next character set, where you should expect a quote-delimited field. Append that to the current result string.
- 4. If the field begins with anything else, then the substring up until the first whitespace character represents an abbreviation key. Locate it and substitute it in. If you don't find the key in the abbrevs dictionary, give a warning and continue on.

5.4 Parsing AUX files

The .aux file contains the filenames of the .bib database file and the .bst style template file, as well as the citations. The get_bibfilenames() method scans through the .aux file and locates a line with \bibdata{...} which contains a filename or a comma-delimited list of filenames, giving the database files. Another line with \bibstyle{...} gives the filename or comma-delimited list of filenames for style templates. The filenames obtained are saved into the filedict attribute – a dictionary whose keys are the file extensions aux, bbl, bib, bst, or tex.

The parse_auxfile() method makes a second pass through the .aux file, this time looking for the citation information. (Auxiliary files are generally quite small, so taking multiple passes through them costs very little time.) Each line with \citation{...} contains a citation key or comma-delimited list of citation keys – each one is added into the citation dictionary (citedict), with a value corresponding to the citation order.

5.5 Parsing BST files

Parsing a .bst file basically involves looking for one of several syntactical structures.

- 1. First, any # present in a line indicates a comment. All text following the # are ignored.
- 2. Any line containing all capital letters and ending in: indicates a section header. The sections recognized are: TEMPLATES, SPECIAL-TEMPLATES, OPTIONS, VARIABLES, and DEFINITIONS. The first three sections (TEMPLATES, SPECIAL-TEMPLATES, and OPTIONS) use template syntax, while the last two (VARIABLES and DEFINITIONS) use Python syntax.
- 3. In the TEMPLATES, SPECIAL-TEMPLATES, or OPTIONS sections of the file, any line ending in an ellipsis (...) means that the following line is a continuation. Thus, the following line is appended to the current one.
- 4. For each var = definition pair found in the VARIABLES section of the file, the code creates a new entry in the user_variables dictionary, with value equal to the given definition.
- 5. For each entrytype = template pair found in the TEMPLATES section of the file, the code creates a corresponding entry in bstdict, with the key given by the entrytype and value given by the template. The code next examines the template definition to see if it contains a nested options block. If so, it adds it to the list of nested templates.
- 6. For each keyword = value pair found in the OPTIONS section of the file, the code creates a new entry in the options dictionary, with the dictionary key being the keyword itself, and the value copied from the right hand side of the option definition.
- 7. For each var = definition pair found in the SPECIAL-TEMPLATES section of the file, the code has to do a little more work than elsewhere. First it creates a new entry in the specials dictionary, with the dictionary key given by the var, and the value given by the definition. It then appens the key to the specials_list. (Since a dictionary is not ordered, we need an order-preserving means of iterating through

the list of specials to make sure that one can always be defined before another that depends on it.) Next it examines the template definition to see if it contains a nested options block. If so, it adds it to the list of nested templates. It also looks to see if there is an ellipsis representing an implicit loop. If so, it adds the template key to the list of "looped templates". Finally it looks to see if the template's key represents an implicitly-indexed variable. If so, it adds the key to the list of implicitly indexed variables.

Once the initial parsing is done, there are several steps in which it analyzes the results:

- 1. Iterating through each of the regular templates, the code looks to see if any of the templates are defined as copies of other templates, as, for example, inbook = incollection. If it finds this kind of definition, then it copies the template from the one (incollection here) to the other (inbook here).
- 2. The code looks at the functions defined in the DEFINITIONS section of the file. If the allow_scripts keyword is set to True, then it goes ahead and evaluates these function definitions so that they will be available during the process of formatting bibliography entries.
- 3. Finally, the code passes each template definition through the validate_templatestr() function to validate that the template has proper syntax.

5.6 Writing the BBL file

Now that all the information is available to Bibulous, we can begin writing the output BBL file. First we write a few lines to the preamble, including the preamble string obtained from the .bib database files. Then, for each citation key we found in the .aux file, we

- 1. Insert any cross-reference data from any other database entries into the current one.
- 2. Define all of the "special variables", including the sortkey and citelabel, as fields within the current entry.

Now that we have all of the sortkeys, we generate the citation_list — the thing we iterate through one by one to format the references in order. At each iteration, we call format_bibitem(), which does the following:

- 1. Write the line \bibitem[citelabel] {citekey} into the .bbl file.
- 2. Import the template corresponding to the current entry's entrytype.
- 3. If there are any user-deefined variables (from the VARIABLES section of the file), then evaluate those variables now, so that they can be used inside the template.
- 4. For each option block in the template, go through and determine how to "simplify" the block. This amounts to locating the first cell in each block that has a defined value, and then replacing the [...] square-bracket-delimited block with its contents. At this point the template variables are still there; only the square brackets have been dropped.
- 5. Now that the optional pieces are all gone, go through each template variable and replace it with the corresponding field from the database entry.
- 6. If there are any nested \textit{...\textit{...} operators in the result, replace odd-level operators with \textup{...} in order to get the right behavior of flipping between italics and regular font.
- 7. If there are any nested \textbf{...}textbf{...} operators in the result, replace odd-level operators with \textup{...} in order to get the right behavior of flipping between bold and regular weight.
- 8. If there are any nested quotation marks in the result, then re-order them according to the American standard. This means having double-quotation-marks at the outermost level, single-quotation-marks inside that, then double inside that, single inside that, and so on. This is messy and difficult code, and so users should always be recommended to use the \enquote(...) LaTeX operator instead of manually-implemented quotation marks.

5.7 Name formatting

One of the more complex tasks needed for parsing BIB files is to resolve the elements of name lists (typically saved in the author and editor fields). In order to know how these should be inserted into a template, it is necessary to know which parts of a given person's name correspond to the first name, the middle name(s), the "prefix" (or "von part"), the last name (or "surname"), and the "suffix" (such as "Jr." or "III"). These five pieces or each person's name are saved as a dictionary, so that a bibliography entry with five authors is represented in <authorlist> as a list of five dictionaries, and each dictionary having keys first, middle, prefix, last, and suffix.

In order to speed up parsing times, the actual mapping of the author or editor fields to authorlist or editorlist is not done until the loop over citation keys performed while writing out the BBL file. The function that product the list-of-dicts parsing result is namestr_to_namedict(namestr).

The default formatting of a namelist into a string to be inserted into the template is performed by format_namelist().

5.7.1 create_namelist()

A BibTeX "name" field can consist of three different formats of names:

- 1. A space-separated list: [firstname middlenames suffix lastname]
- 2. A two-element comma-separated list: [prefix lastname, firstname middlenames]
- 3. A three-element comma-separated list: [prefix lastname, suffix, firstname middlenames]

So, an easy way to separate these three categories is by counting the number of commas that appear. The trickiest part here is that although we can use and as a name separator, we are only allowed to do so if and occurs at the top brace level.

In addition, in order to make name parsing more flexible for nonstandard names, Bibulous adds two more name formats to this list:

- 4. A four-element comma-separated list: [firstname, middlenames, prefix, lastname]
- 5. A five-element comma-separated list: [firstname, middlenames, prefix, lastname, suffix]

For each name in the field, we parse the name tokens into a dictionary. We then compile all of the dictionaries into a list, ordered by the appearance of the names in the input field.

5.7.2 format_namelist()

Given a namelist (list of dictionaries), we glue the name elements together into a single string, incorporating all of the format options selected by the user in the template file. This includes calls to namedict_to_formatted_namestr(), and to initialize_name() if converting any name tokens to initials.

5.8 Generating sortkeys

If the user's style template file selects the citation order to be citenum or none, then creating the ordered citation list is as simple as listing the citation keys in order of their citation appearance, which was recorded as the value in the citation dictionary. If the user instead chooses the citation order to be citekey, then all that is needed is to sort the citation keys alphabetically. Similar operations follow for the various citation order options, but the difficult lies in correctly sorting in the presence of non-ASCII languages, and especially in the presence of LaTeX markup of non-ASCII names. For a citation sorting order that requires using author names, any LaTeX markup needs to be converted to its Unicode equivalent prior to sorting. Using unicode allows the sorting to be done with any input languages, and allows the sorting order to be locale-dependent.

create_citation_list() is the highest-level function for generating the citation list. For each citation key, it calls <code>generate_sortkey()</code>, which is the workhorse function for including all of the various options when generating the key to use for sorting the list. A key part of the function is a call to <code>purify_string()</code>, which removes unnecessary LaTeX markup elements and then calls <code>latex_to_utf8()</code> to convert LaTeX-markup non-ASCII characters to Unicode. It is only after all of these conversions that the final sorting is performed and the sorted citation list returned.

5.9 Testing

The suite of regression tests for Bibulous consist of various template definitions and database entries designed to test individual features of the program. The basic approach of the tests is as follows:

1. Once a change is made to the code (to fix a bug or add functionality), the developer also adds an entry to the test/test1.bib file, where the entry's "entrytype" is named in such a way to give an indication of what the test is for. For example, the entry in the BIB file may be defined with:

```
@initialize1{...
```

where the developer provides an author field in the entry where one or more authors have names which are difficult to for generating initials correctly. The developer should also include at least a 1-line comment about the purpose of the entry as well. To make everything easy to find, use the entrytype as the entry's key as well. Thus, the example above would use:

```
@initialize1{initialize1, ...
```

- 2. If the above new entry is something which can be checked with normal options settings, then the developer should add a corresponding line in the BST file defining how that new entrytype (i.e. initialize1) should be formatted. If *different* options settings are needed, then a new BST file is needed. Only a minimalist file is generally needed: the file can, for example, contain one line defining a new entrytype and one line to define the new option setting. You can define all of the other options if you want, but these are redundant and introduce a number of unnecessary "overwriting option value..." warning messages.
- 3. Next, the developer should add a line \citation{entrytype} to the AUX file where the entrytype is the key given in the new entry of the BIB file you just put in (e.g. initialize1). This is the same as the entrytype to keep everything consistent.
- 4. Next, the developer needs to add two lines to the test1_target.bbl file to say what the formatted result should look like. Take a look at other lines to get a feel for how these should look, and take in consideration the form of the template just added to the BST file.
- 5. Finally, run bibulous_test.py to check the result. This script will load the modified BIB and BST files and will write out several formatted BBL file test1.bbl etc. It will then run a diff program on the output file versus the target BBL file to see if there are any differences between the target and actual output BBL files.

5.10 Generating the documentation

The documentation is written in reStructuredText (RST) and converted to HTML using Sphinx. Sphinx can also use LaTeX to convert the HTML files into a PDF.

From the bibulous repository doc/subfolder, run make html to generate the HTML documentation. The result can be found in doc/_build/html/, with index.html as the main file. To generate the PDF documentation, run make latexpdf from the doc/subfolder, with the result found at doc/_build/latex/Bibulous.pdf.

While the documentation is saved in the doc/ folder on the main branch, this is not automatically converted into viewable, linked HTML on GitHub. To achieve that requires pushing the updated docs into the gh-pages branch. One way of doing this is the following. Make a local copy of the main branch's doc/_build/html/ folder. Switch

to the gh-pages branch (i.e. git checkout gh-pages) and replace everything there with the locally-copied doc/_build/html/ folder contents. Then update: git add -A and git push origin gh-pages. And switch back to the main branch, git checkout master.

5.11 Updating the PyPI package

From the bibulous base folder, run:

```
python setup.py sdist --formats=gztar,zip
```

to create the package locally, and then run:

```
python setup.py sdist upload
```

to update the PIP package online.

5.12 Miscellaneous notes

The code includes two different variables, citekey and entrykey which for any given entry are always identical. So it would appear that they are redundant. But the keys in the citedict dictionary, and the keys specifying each entry in the database, belong to different sets. That is, the list of entry keys can be from every entry in the database, even entries that were not cited. The list of citation keys, however, contains only those keys that were cited, and so can be a much smaller list.

EXAMPLES

6.1 Example 1

The following example is taken from a question posted at http://tex.stackexchange.com/questions/147675/bibtex-scientific-style-with-pages-at-the-end, where the desired bibliography format is given as:

```
Journals

1. Holzwarth G., Eckart G. Fluid-dynamical approximation for finite Fermi systems // Nucl. Phys. – 1979. – Vol. A325. P. 1 – 30.

Books

2. Bertsch G. F., Broglia R. A. Oscillations in finite quantum systems. Ch. 6. – Cambridge: Cambridge University, 1994. – 150 p.

Chapters

3. Van der Woude A. The electric giant resonances // Electric and magnetic giant resonances in nuclei / Ed. by J. Speth. – Singapure: World Scientific P.C., 1991. – P. 99 – 232.

Conference or symposium proceedings

4. Smolanzuk R., Skalski J., Sobiczewski A. Masses and half-life of superheavy elements // Proc. of the International Workshop 24 on Gross Properties of Nuclei and Nuclear Excitations / Ed. by H. Feldmeier et al. – GSI, Darmstadt, 1996. – P. 35 – 42.
```

With Bibulous, we can easily provide templates that provide the formatting that the OP asks for:

```
TEMPLATES:
article = \textit{<au>} <title> // <journal> -- <year>. -- Vol.~<volume>. P.~[
book = \textit{<au>} <title>. [Ch.~<chapter>. ]-- <address>: <publisher>, <year>.[ --

<startpage>~p.]
inbook = \textit{<au>} <title> // <booktitle>[ / Ed. by <ed.if_singular(editorlist,_</pre>
→nothing, etal_message)>]. -- <address>: ...
         <publisher>, <year>.[ -- P.~<startpage> -- <endpage>| -- P.~<startpage>| --_
inproceedings = \textit{<au>} <title> // <booktitle>[ / Ed. by <ed.if_</pre>
→singular(editorlist, nothing, etal_message)>]. -- ...
                <publisher>, <address>, <year>. -- P.~[<startpage> -- <endpage>|

<startpage>|<eid>|].
SPECIAL-TEMPLATES:
authorlist = <author.to_namelist()>
editorlist = <editor.to_namelist()>
\verb|authorname.n| = [\langle authorlist.n.prefix \rangle] \langle authorlist.n.last \rangle [\langle authorlist.n.first.]
→initial()>.][ <authorlist.n.middle.initial()>.][, <authorlist.n.suffix>.]
au = <authorname.0>, ..., <authorname.9>
ed = [<editorlist.0.first.initial()>. ][ <editorlist.0.middle.initial()>. ][
→ <editorlist.0.prefix> ] <editorlist.0.last>[, <editorlist.0.suffix>.]
OPTIONS:
nothing = {}
```

so that if we use this template together with the following database file:

```
@article { Holzwarth,
    author = {G. Holzwarth and G. Eckart},
    title = {Fluid-dynamical approximation for finite Fermi systems},
    journal = {Nucl. Phys.},
    year = 1979,
    volume = \{A325\},
   pages = \{1-30\}
@book { Bertsch,
   author = {G. F. Bertsch and R. A. Broglia},
   title = {Oscillations in finite quantum systems},
   chapter = 6,
   address = {Cambridge},
   publisher = {Cambridge University},
   year = 1994,
    pages = 150
@inbook { Woude,
   author = {A. {Van der Woude}},
   title = {The electric giant resonances},
   booktitle = {Electric and magnetic giant resonances in nuclei},
   editor = {J. Speth},
   address = {Singapore},
   publisher = {World Scientific P.C.},
   year = 1991,
   pages = \{99-232\}
@inproceedings { Smolanzuk,
   author = {R. Smolanzuk and J. Skalski and A. Sobiczewski},
    title = {Masses and half-life of superheavy elements},
   booktitle = {Proc.\ of the International Workshop 24 on Gross Properties of...
→ Nuclei and Nuclear Excitations },
    editor = {H. Feldmeier and others},
   address = {Darmstadt},
   publisher = {GSI},
   year = 1996,
   pages = {35-42}
```

then we get the formatted result shown below

References

- 1. Holzwarth G., Eckart G. Fluid-dynamical approximation for finite Fermi systems // Nucl. Phys. 1979. Vol. A325. P. 1 30.
- Bertsch G. F., Broglia R. A. Oscillations in finite quantum systems. Ch. 6. Cambridge: Cambridge University, 1994. – 150 p.
- 3. Van der Woude A. The electric giant resonances // Electric and magnetic giant resonances in nuclei / Ed. by J. Speth. Singapore: World Scientific P.C., 1991. P. 99 232.
- Smolanzuk R., Skalski J., Sobiczewski A. Masses and half-life of superheavy elements // Proc. of the International Workshop 24 on Gross Properties of Nuclei and Nuclear Excitations / Ed. by H. Feldmeier, et al. GSI, Darmstadt, 1996. P. 35 42.

6.2 Example 2

The next example is taken from the bibliography style found in: Dimitri Mihalas and James Binney, _Galactic Astronomy: Structure and Kinematics_, 2nd ed. (W. H. Freeman, New York, 1981). A snapshot from the book's bibliography looks like

- (B12) Blaauw, A. and Schmidt, M. (eds.). 1965. *Galactic Structure*. (Chicago: University of Chicago Press).
- (B13) Bok, B. J. 1977. Pub. Astron. Soc. Pacific. 89:597.
- (B14) Bosma, A. 1978. Ph.D. Thesis, University of Groningen, Netherlands.
- (B15) Burke, B. F. 1957. Astron. J. 62:90.
- (B16) Burton, W. B. 1970. Astron. and Astrophys. 10:76.
- (B17) Burton, W. B. 1972. Astron. and Astrophys. 19:51.
- (B18) Burton, W. B. 1976. Ann. Rev. Astron. and Astrophys. 14:275.
- (B19) Burton, W. B. and Gordon, M. A. 1978. Astron. and Astrophys. 63:7.
- (C1) Chiu, H.-Y. and Muriel, A. (eds.). 1970. Galactic Astronomy. (New York Gordon and Breach).
- (C2) Cohen, R. J. and Davies, R. D. 1976. Mon. Not. Roy. Astron. Soc. 175:1.
- (D1) Dickman, R. L. 1978. Astrophys. J. Supp. 37:407.
- (E1) Emerson, D. T. 1978. Astron. and Astrophys. 63:L29.
- (F1) Fichtel, C. E. and Stecker, F. W. (eds.). 1977. The Structure and Content of the Galaxy and Galactic Gamma Rays (Washington, D.C.: NASA).
- (F2) Fletcher, E. 1963. Astron. J. 68:407.

To produce this style, we can define the following templates:

```
TEMPLATES:
article = <au> <year>. \textit{<journal>} \textbf{<volume>}:<pages>.
book = [<au>|<ed.if_singular(editorlist, eds_msg1, eds_msg2)>|]. <year>. \textit{
→<title>}. (<address>: <publisher>).
proceedings = book
phdthesis = <au> <year>. Ph.D. Thesis, <university>.
SPECIAL-TEMPLATES:
authorlist = <author.to_namelist()>
editorlist = <editor.to_namelist()>
authorname.n = [<authorlist.n.prefix> ]<authorlist.n.last>[, <authorlist.n.first.</pre>
→initial()>.][ <authorlist.n.middle.initial()>.][,
                                                     <authorlist.n.suffix>.]
au = <authorname.0>, ..., { and } <authorname.9>
editorname.n = [<editorlist.n.prefix> ]<editorlist.n.last>[, <editorlist.n.first.</pre>
→initial()>.][ <editorlist.n.middle.initial()>.][, <editorlist.n.suffix>.]
ed = <editorname.0>, ...,{ and }<editorname.5>
authorlabel = [<authorlist.0.prefix>|<authorlist.0.last>]
editorlabel = [<editorlist.0.prefix>|<editorlist.0.last>]
sortkey = [<authorlabel><year.zfill(4)>|<editorlabel><year>]
citelabel = (<citealnum>)
OPTIONS:
eds_msg1 = { } (ed.)
eds_msq2 = { } (eds.)
use_citeextract = True
```

so that with the following database file:

```
@book{Blaauw1965,
   editor = {A. Blaauw and M. Schmidt},
   title = {Galactic Structure},
   address = {Chicago},
   publisher = {University of Chicago Press},
```

6.2. Example 2 37

```
year = 1965
@article{Bok1977,
   author = \{B. J. Bok\},
    journal = {Publ. Astron. Soc. Pacific},
    year = 1977,
    volume = 89,
   pages = 597
@phdthesis(Bosma1978,
   author = {A. Bosma},
   university = {University of Groningen, Netherlands},
   year = 1978
@article { Burke1957,
   author = {B. F. Burke},
    journal = {Astron. J.},
   year = 1957,
   volume = 62,
   pages = 90
@article{Burton1970,
   author = {W. B. Burton},
   journal = {Astron. and Astrophys.},
   year = 1970,
   volume = 10,
   pages = 76
@article{Burton1972,
   author = {W. B. Burton},
   journal = {Astron. and Astrophys.},
    year = 1972,
    volume = 19,
   pages = 51
@article{Burton1976,
   author = {W. B. Burton},
   journal = {Ann. Rev. Astron. and Astrophys.},
   year = 1976,
   volume = 14,
   pages = 275
@article{Burton1978,
   author = {W. B. Burton and M. A. Gordon},
    journal = {Astron. and Astrophys.},
   year = 1978,
   volume = 63,
   pages = 7
@book { Chiu1970,
   editor = {H.-Y. Chiu and A. Muriel},
   title = {Galactic Astronomy},
   year = 1970,
   address = {New York},
   publisher = {Gordon and Breach}
@article{Cohen1976,
```

```
author = {R. J. Cohen and R. D. Davies},
    journal = {Mon. Not. Roy. Astron. Soc.},
   year = 1976,
    volume = 175,
    pages = 1
@article { Dickman1978,
   author = {R. L. Dickman},
    journal = {Astrophys. J. Supp.},
   year = 1978,
    volume = 37,
   pages = 407
@article { Emerson1978,
   author = {D. T. Emerson},
    journal = {Astron. and Astrophys.},
   year = 1978,
    volume = 63,
   pages = \{L29\},
@proceedings(Fichtel1977,
    editor = {C. E. Fichtel and F. W. Stecker},
   title = {The Structure and Content of the Galaxy and Galactic Gamma Rays},
   address = {Washington, D.C.},
   publisher = {NASA},
   year = 1977
```

we get the following formatted result

References

```
(B1) Blaauw, A. and Schmidt, M. (eds.). 1965. Galactic Structure. (Chicago: University of Chicago Press).
(B2) Bok, B. J. 1977. Publ. Astron. Soc. Pacific 89:597.
(B3) Bosma, A. 1978. Ph.D. Thesis, University of Groningen, Netherlands.
(B4) Burke, B. F. 1957. Astron. J. 62:90.
(B5) Burton, W. B. 1970. Astron. and Astrophys. 10:76.
(B6) Burton, W. B. 1972. Astron. and Astrophys. 19:51.
(B7) Burton, W. B. 1976. Ann. Rev. Astron. and Astrophys. 14:275.
(B8) Burton, W. B. and Gordon, M. A. 1978. Astron. and Astrophys. 63:7.
(C1) Chiu, H.-Y. and Muriel, A. (eds.). 1970. Galactic Astronomy. (New York: Gordon and Breach).
(C2) Cohen, R. J. and Davies, R. D. 1976. Mon. Not. Roy. Astron. Soc. 175:1.
(D1) Dickman, R. L. 1978. Astron. and Astrophys. 63:L29.
(F1) Fichtel, C. E. and Stecker, F. W. (eds.). 1977. The Structure and Content of the Galaxy and Galactic Gamma Rays. (Washington, D.C.: NASA).
```

6.3 Example 3

The next example illustrates the style used by the Optical Society of America (OSA) for their journals. The following formatted bibliography

6.3. Example 3 39

References

- D. Jacquemart and L. S. Rothman, "The next HITRAN edition: description of new parameters and formats", poster (2012).
- 2. B. E. Baver, "Color imaging array," USA patent 3971065 (1976).
- 3. S. Boyd and L. Vandenberghe, "Convex optimization class slides", presentation (2006).
- J. Bohannon, R. Goldstein, and A. Herschkowitsch, "Can people distinguish pâté from dog food?" American Association of Wine Economists, Tech. Rep. 51757 (2009).
- 5. M. R. Descour, ed., Imaging Spectrometry VII, vol. 4480 (2002).
- 6. CGAL Project, CGAL user and reference manual (CGAL Project, 2006).
- 7. D. E. Knuth, "Convolution polynomials" (1992). (url: http://arxiv.org/pdf/math/9207221.pdf)
- E. R. Tkaczyk, "Vectorial laws of refraction and reflection using the cross product and dot product," Opt. Lett. 37, 972–974 (2012).
- M. R. Descour and E. L. Dereniak, "Nonscanning no-moving-parts imaging spectrometer," in *Imaging Spectrometry*, in Proceedings of SPIE 2480, pp. 48–64 (1995).
- 10. www.zemax.com
- 11. W. J. Smith, Modern Lens Design: A Resource Manual (McGraw-Hill, 1992).
- R. P. Feynman, "Negative probability," in Quantum Implications Essays in Honour of David Bohm, B. J. Hiley and F. D. Peat, eds, pp. 235–248 (Routledge & Kegan Paul, 1987).
- 13. M. P. Rimmer, "Optical aberration coefficients," M.S. dissertation (University of Rochester, 1963).
- 14. L.-V. de Broglie, "On the theory of quanta," Ph.D. dissertation (University of Paris, 1925). A translation by A. F. Kracklauer of the original, Recherches sur la théorie des quanta, (Ann. de Phys., 10th series, v. III, Jan-Feb 1925).
- M. Silverman, "Optical detection theory," in Probing the Atom: Interactions of Coupled States, Fast Beams, and Loose Electrons, Chap. 6, pp. 139–159 (Princeton University Press, 1998).

is obtained using the template

```
TEMPLATES:
article = \langle au \rangle, \equote{[\href{\langle url \rangle} {\langle title \rangle} | \href{\langle doi \rangle} {\langle title \rangle} | \langle title \rangle],}
→<journal> \textbf{<volume>}, { }...
        [<startpage>--<endpage>|<startpage>|<eid>|] (<year>).[ <note>]
book = [<au>|<ed>|<organization>|<institution>|], \textit{<title>}[, <edition.</pre>
→ordinal()>~ed.][, vol.~<volume>]{ }...
        (<publisher>, <year>)[ Chap.~<chapter>][, pp.~<startpage>--<endpage>|, p.~
→<startpage>].[ <note>]
endnote = <note>
inbook = incollection
incollection = <au>, \enquote{<title>,} in \textit{<booktitle>}, [vol.~<volume>, ][
→<edition.ordinal()>~ed., ]...
        [<ed>, ][<series>, ][Chap.~<chapter>, ][pp.~<startpage>--<endpage>|p.~

</startpage>|<eid>|] (<publisher>, <year>).[ <note>]

inproceedings = <au>, \enquote{<title>,} in \textit{<booktitle>}[, <ed>][, in <series>
→][ \textbf{<volume>}]...
        [, pp.~<startpage>--<endpage>|, p.~<startpage>|, <eid>] (<year>).[ <note>]
[, <edition.ordinal()>~ed.|, ver.~<version>][, <ed>] (<organization>, <year>
→)[, Chap.~<chapter>]...
        [, pp.~<startpage>--<endpage>|, p.~<startpage>].[ <note>]
mastersthesis = <au>, \enquote{<title>,} M.S. dissertation (<school>, <year>).[ <note>
\hookrightarrow
misc = [<au>, ][\enquote{<title>}][ (<organization>, <year>)| (<year>)].[ <note>][_
→ (url: \url{<url>})]
patent = <au>, \enquote{<title>,} <nationality> patent <number> (<year>).[ <note>]
phdthesis = <au>, \enquote{<title>,} Ph.D. dissertation (<school>, <year>).[ <note>]
poster = <au>, \enquote{<title>}, poster [presented at <conference>]([<month> ]<year>
\rightarrow).[ <note>]
presentation = <au>, \enquote{<title>}, presentation [presented at <conference>]([
proceedings = <ed>, \textit{<title>}, [vol.~<volume>][, <edition.ordinal()>~ed.] (
\rightarrow<year>).[ <note>]
techreport = [<au>|<ed>|<organization>|<institution>|], \enquote{<title>,}
→<institution>, Tech. Rep. <number> (<year>).[ <note>]
website = \href{[<url>|<doi>]}{[<url>|<doi>]}
```

```
SPECIAL-TEMPLATES:
sortkey = <citenum.zfill(2)>
                                         ## order the reference list in the same_
→order as the citations were given in
                                          ## use numeric reference labels
citelabel = <citenum>.
OPTIONS:
maxauthors = 25
                                          ## if more than 25 authors, then use "et al.
→" format (only applied when using the default "<au>" variable)
minauthors = 9
                                         ## if "maxauthors" is exceeded, then_
→display only the first "minauthors" and et al. (only applied when using the default
→"<au>" variable)
maxeditors = 5
                                          ## if more than 5 editors, then use "et al.
→ " format (only applied when using the default "<ed>" variable)
mineditors = 5
                                          ## if "maxeditors" is exceeded, then...
→display only the first "mineditors" and et al. (only applied when using the default
→"<ed>" variable)
procspie_as_journal = True
                                          ## treat SPIE conference proceedings like a_
→journal, with journal name 'Proc. SPIE'
sort_case = False
                                          ## whether to use case-sensitive citation...
→sorting (default is False)
use_abbrevs = True
                                          ## whether to use the abbreviation strings...
→ (mainly useful for debugging)
french_initials = False
                                          ## whether to use French initials (i.e._
→Jean-Paul -> J.-P. and not J., also Philippe -> Ph. not P.)
bibitemsep = 0pt
                                          ## the extra vertical separation between...
→each bibliography entry. Defining this option *overrides* LaTeX's existing...
→definition for \bibitemsep.
etal_message = , \textit{et~al.}
                                         ## what text to place at the end of an_
\rightarrowauthorlist when there are too many
edmsg1 = , ed.
                                          ## the label to use after the editor...
→namelist when only one editor is present
edmsg2 = , eds
                                          ## the label to use after the editor...
→namelist when multiple editors are present
## The remaining exceptions below are really for users rather than for the journal. _
→Maybe put these someplace else?
undefstr = ???
                                          ## what string to replace undefined.
→required fields with
replace_newlines = True
                                         ## whether to replace newline characters...
→inside fields with spaces
allow_scripts = False
                                         ## whether to allow user scripts in BST_
-→files
case_sensitive_field_names = False ## whether to allow field names to be case...
→sensitive (i.e. so that "BOOK" will not be the same entrytype as "book")
use citeextract = True
                                          ## whether to automatically perform.
⇒citation extraction each time Bibulous is run
name separator = and
                                         ## the string used to separate names in a...
→BibTeX-format field of names
VARIABLES:
DEFINITIONS:
```

and the database file

6.3. Example 3 41

```
@String{ol
                = {Opt. Lett.}}
@String{procspie = {Proc. SPIE}}
@Article { Tkaczyk2012,
 title = {Vectorial laws of refraction and reflection using the cross product and,
→dot product },
 author = {Eric R. Tkaczyk},
  journal = ol,
 volume = \{37\},
 year = \{2012\},
 pages = \{972 - 974\},
 doi = \{http://dx.doi.org/10.1364/OL.37.000972\}
@Book { Smith-1992,
 title = {Modern Lens Design: A Resource Manual},
 year = \{1992\},
 address = {New York},
 author = {Warren J. Smith},
 publisher = {McGraw-Hill}
@Inbook{Silverman1998,
 title = {Optical detection theory},
 author = {Mark Silverman},
  pages = \{139 - -159\},
  publisher = {Princeton University Press},
 year = \{1998\},
 booktitle = {Probing the Atom: Interactions of Coupled States, Fast Beams, and
→Loose Electrons},
 chapter = \{6\}
@Incollection {Feynman1987,
 title = {Negative probability},
  author = {Richard P. Feynman},
 booktitle = {Quantum Implications --- Essays in Honour of {D}avid Bohm},
  publisher = {Routledge \& Kegan Paul},
  year = \{1987\},
 pages = \{235 - 248\},
  editor = {B. J. Hiley and F. D. Peat}
@Inproceedings {Descour1995a,
 title = {Nonscanning no-moving-parts imaging spectrometer},
 author = {Michael R. Descour and Eustace L. Dereniak},
 booktitle = {Imaging Spectrometry},
 year = \{1995\},
 pages = \{48 - -64\},
 series = procspie,
 volume = {2480},
  crossref = {ProcSPIE2480}
@Manual { CGAL2006,
 title = {CGAL user and reference manual},
 year = \{2006\},
```

```
month = {July},
  organization = {CGAL Project}
@Mastersthesis { RimmerDiss,
 title = {Optical aberration coefficients},
 author = {Matthew Peter Rimmer},
 school = {University of Rochester},
 year = \{1963\}
@Misc{Knuth1992,
 title = {Convolution polynomials},
 author = {Donald E. Knuth},
 vear = \{1992\},
 url = {http://arxiv.org/pdf/math/9207221.pdf}
@Patent {Bayer1976,
  title = {Color imaging array},
  author = {Bryce E. Bayer},
 number = \{3971065\},
 year = \{1976\},
 yearfiled = \{1975\},
 assignee = {Eastman Kodak Company},
 nationality = {USA}
@Phdthesis { BroglieDiss,
 title = {On the theory of quanta},
 author = {Louis-Victor de Broglie},
  school = {University of Paris},
 year = \{1925\},
 note = {A translation by A. F. Kracklauer of the original, Recherches sur la th\{\'e\}
→orie des quanta, (Ann. de Phys., 10th series, v. III, Jan-Feb 1925).}
@Poster{Jacquemart2012,
 title = {The next HITRAN edition: description of new parameters and formats},
 author = {David Jacquemart and Laurence S. Rothman},
 address = {[location unknown]},
 year = \{2012\}
@Presentation{Boyd2006,
 title = {Convex optimization - class slides},
  author = {Stephen Boyd and Lieven Vandenberghe},
 year = \{2006\}
@Proceedings {ProcSPIE4480,
 title = {Imaging Spectrometry VII},
 year = \{2002\},
 editor = {Michael R. Descour},
 series = procspie,
  volume = {4480}
```

(continues on next page)

6.3. Example 3 43

```
@Techreport{Bohannon2009,
    title = {Can people distinguish p{\^a}t{\'e} from dog food?},
    author = {John Bohannon and Robin Goldstein and Alexis Herschkowitsch},
    institution = {American Association of Wine Economists},
    year = {2009},
    number = {51757},
    month = {April}
}

@Website{Zemax,
    url = {www.zemax.com},
    organization = {ZEMAX Development Corp.}
}
```

6.4 Example 4

The next example illustrates the style used by the society SPIE for their journals. The following formatted bibliography

References

- D. Jacquemart and L. S. Rothman, "The next HITRAN edition: description of new parameters and formats", poster (2012).
- 2 B. E. Bayer, "Color imaging array," USA patent 3971065 (1976).
- 3 S. Boyd and L. Vandenberghe, "Convex optimization class slides", presentation (2006).
- 4 J. Bohannon, R. Goldstein, and A. Herschkowitsch, "Can people distinguish pâté from dog food?" American Association of Wine Economists, 51757 (2009).
- 5 M. R. Descour, ed., Imaging Spectrometry VII, vol. 4480 (2002).
- 6 CGAL Project, CGAL user and reference manual (CGAL Project, 2006)
- 7 D. E. Knuth, "Convolution polynomials" (1992). (url: http://arxiv.org/pdf/math/9207221.pdf)
- E. R. Tkaczyk, "Vectorial laws of refraction and reflection using the cross product and dot product," Opt. Lett. 37(5), 972–974 (2012).
- 9 M. R. Descour and E. L. Dereniak, "Nonscanning no-moving-parts imaging spectrometer," in *Imaging Spectrometry*, in Proc. SPIE 2480, pp. 48–64 (1995).
- 10 www.zemax.com
- 11 W. J. Smith, Modern Lens Design: A Resource Manual McGraw-Hill, New York (1992).
- 12 R. P. Feynman, "Negative probability," in Quantum Implications Essays in Honour of David Bohm, B. J. Hiley and F. D. Peat, eds, pp. 235–248 (Routledge & Kegan Paul, 1987).
- 13 M. P. Rimmer, "Optical aberration coefficients," M.S. dissertation (University of Rochester, 1963).
- 14 L.-V. de Broglie, "On the theory of quanta," Ph.D. dissertation (University of Paris, 1925). A translation by A. F. Kracklauer of the original, Recherches sur la théorie des quanta, (Ann. de Phys., 10th series, v. III, Jan-Feb 1925).
- 15 M. Silverman, "Optical detection theory," in Probing the Atom: Interactions of Coupled States, Fast Beams, and Loose Electrons, Chap. 6, pp. 139–159 (Princeton University Press, 1998).

is obtained using the template

```
inproceedings = <au>, \enguete{<title>,} in \textit{<booktitle>}[, <ed>][, in <series>
→][ \textbf{<volume>}]...
        [, pp.~<startpage>--<endpage>|, p.~<startpage>|, <eid>] (<year>).[ <note>]
manual = [<au>|<ed>|<organization>|<institution>|], \textit{<title>}[, vol.~<volume>].
        [, <edition.ordinal()>~ed.|, ver.~<version>][, <ed>] (<organization>, <year>
→) [, Chap.~<chapter>]...
        [, pp.~<startpage>--<endpage>|, p.~<startpage>].[ <note>]
mastersthesis = <au>, \enquote{<title>,} M.S. dissertation (<school>, <year>).[ <note>
\hookrightarrow
misc = [<au>, ][\enquote{<title>}][ (<organization>, <year>)| (<year>)].[ <note>][_

    (url: \url{<url>})]
patent = <au>, \enquote{<title>,} <nationality> patent <number> (<year>).[ <note>]
phdthesis = <au>, \enquote{<title>,} Ph.D. dissertation (<school>, <year>).[ <note>]
poster = <au>, \enquote{<title>}, poster [presented at <conference>]([<month> ]<year>
→).[ <note>]
presentation = <au>, \enquote{<title>}, presentation [presented at <conference>]([
→<month> ]<year>).[ <note>]
proceedings = <ed>, \textit{<title>}, [vol.~<volume>][, <edition.ordinal()>~ed.] (
techreport = [<au>|<ed>|<organization>|<institution>|], \enguote{<title>,}
→<institution>, <number> (<year>).[ <note>]
website = \href{[<url>|<doi>]}{[<url>|<doi>]}
SPECIAL-TEMPLATES:
sortkey = <citenum.zfill(2)>
                                         ## order the reference list in the same.
→order as the citations were given in
citelabel = <citenum>
                                          ## use numeric reference labels
OPTIONS:
maxauthors = 3
                                          ## if more than 3 authors, then use "et al.
→" format (only applied when using the default "<au>" variable)
minauthors = 1
                                          ## if "maxauthors" is exceeded, then.
→display only the first "minauthors" and et al. (only applied when using the default
→"<au>" variable)
maxeditors = 3
                                          ## if more than 3 editors, then use "et al.
→ " format (only applied when using the default "<ed>" variable)
                                          ## if "maxeditors" is exceeded, then...
mineditors = 1
→display only the first "mineditors" and et al. (only applied when using the default
→"<ed>" variable)
procspie_as_journal = False
                                          ## treat SPIE conference proceedings like a...
→ journal, with journal name 'Proc. SPIE'
sort_case = False
                                          ## whether to use case-sensitive citation.
→sorting (default is False)
use_abbrevs = True
                                          ## whether to use the abbreviation strings_
→ (mainly useful for debugging)
bibitemsep = 0pt
                                          ## the extra vertical separation between.
→each bibliography entry. Defining this option *overrides* LaTeX's existing.
\rightarrowdefinition for \bibitemsep.
etal_message = , \textit{et~al.}
                                         ## what text to place at the end of an_
→authorlist when there are too many
edmsq1 = , ed.
                                          ## the label to use after the editor_
→namelist when only one editor is present
edmsq2 = , eds
                                          ## the label to use after the editor...
→namelist when multiple editors are present
## The remaining exceptions below are really for users rather than for the journal.
                                                                         (continues on next page)
→Maybe put these someplace else?
```

6.4. Example 4 45

```
undefstr = ???
                                             ## what string to replace undefined_
→required fields with
replace_newlines = True
                                             ## whether to replace newline characters_
\rightarrowinside fields with spaces
allow_scripts = False
                                             ## whether to allow user scripts in BST_
→files
case_sensitive_field_names = False
                                             ## whether to allow field names to be case...
\rightarrowsensitive (i.e. "BOOK" will not be the same entrytype as "book")
name_separator = and
                                            ## the string used to separate names in a_
\hookrightarrowBibTeX-format field of names
VARIABLES:
DEFINITIONS:
```

and the database file

```
@String{ol = {Opt. Lett.}}
@String{procspie = {Proc. SPIE}}
@Article { Tkaczyk2012,
 title = {Vectorial laws of refraction and reflection using the cross product and,
→dot product},
 author = {Eric R. Tkaczyk},
 journal = ol,
 volume = \{37\},
 number = \{5\},
 year = \{2012\},
 pages = \{972 - 974\},
 doi = \{http://dx.doi.org/10.1364/OL.37.000972\}
@Book { Smith-1992,
 title = {Modern Lens Design: A Resource Manual},
 year = \{1992\},
 address = {New York},
 author = {Warren J. Smith},
 publisher = {McGraw-Hill}
@Inbook{Silverman1998,
 title = {Optical detection theory},
 author = {Mark Silverman},
 pages = \{139 - -159\},
  publisher = {Princeton University Press},
 year = \{1998\},
 booktitle = {Probing the Atom: Interactions of Coupled States, Fast Beams, and
→Loose Electrons},
 chapter = \{6\}
@Incollection {Feynman1987,
 title = {Negative probability},
 author = {Richard P. Feynman},
 booktitle = {Quantum Implications --- Essays in Honour of {D}avid Bohm},
 publisher = {Routledge \& Kegan Paul},
 year = \{1987\},
```

47

```
pages = \{235 - 248\},
  editor = {B. J. Hiley and F. D. Peat}
@Inproceedings (Descour1995a,
 title = {Nonscanning no-moving-parts imaging spectrometer},
  author = {Michael R. Descour and Eustace L. Dereniak},
  booktitle = {Imaging Spectrometry},
 year = \{1995\},
  pages = \{48 - - 64\},
  series = procspie,
 volume = {2480},
  crossref = {ProcSPIE2480}
@Manual { CGAL2006,
 title = {CGAL user and reference manual},
  year = \{2006\},\
 month = {July},
  organization = {CGAL Project}
@Mastersthesis { RimmerDiss,
 title = {Optical aberration coefficients},
 author = {Matthew Peter Rimmer},
 school = {University of Rochester},
 year = \{1963\}
@Misc{Knuth1992,
 title = {Convolution polynomials},
  author = {Donald E. Knuth},
 year = \{1992\},
 url = {http://arxiv.org/pdf/math/9207221.pdf}
@Patent {Bayer1976,
 title = {Color imaging array},
  author = {Bryce E. Bayer},
 number = \{3971065\},
 year = \{1976\},
 yearfiled = \{1975\},
 assignee = {Eastman Kodak Company},
 nationality = {USA}
@Phdthesis { BroglieDiss,
 title = {On the theory of quanta},
 author = {Louis-Victor de Broglie},
 school = {University of Paris},
 year = \{1925\},
 note = {A translation by A. F. Kracklauer of the original, Recherches sur la th{\'e}
→orie des quanta, (Ann. de Phys., 10th series, v. III, Jan-Feb 1925).}
@Poster{Jacquemart2012,
 title = {The next HITRAN edition: description of new parameters and formats},
                                                                            (continues on next page)
```

6.4. Example 4

```
author = {David Jacquemart and Laurence S. Rothman},
  address = {[location unknown]},
  year = \{2012\}
@Presentation { Boyd2006,
  title = {Convex optimization - class slides},
  author = {Stephen Boyd and Lieven Vandenberghe},
  year = \{2006\}
@Proceedings {ProcSPIE4480,
  title = {Imaging Spectrometry VII},
  year = \{2002\},
  editor = {Michael R. Descour},
  series = procspie,
  volume = {4480}
@Techreport {Bohannon2009,
  title = {Can people distinguish p\{\^a\}t\{\'e\} from dog food?},
  author = {John Bohannon and Robin Goldstein and Alexis Herschkowitsch},
  institution = {American Association of Wine Economists},
  year = \{2009\},
  number = \{51757\},
  month = {April}
@Website { Zemax,
  url = {www.zemax.com},
  organization = {ZEMAX Development Corp.}
```

6.5 Example 5

The following example is taken from a question posted at http://tex.stackexchange.com/questions/160737/bold-labels-and-more-with-custom-bibtex-bst-and-author-year, where an answer to the original poster's question is given as

Literatur

```
Kornmeier, M.: Wissenschaftlich schreiben leicht gemacht: für Bachelor, Master und Dissertation, 4. Aufl., (UTB; 3154), Bern u.a.: Haupt 2011.
Nicol, N.; Albrecht, R.: Wissenschaftliche Arbeiten schreiben mit Word 2010, 7. Aufl., München u.a.: Addison Wesley 2011.
Scholz, S.; Menzl, S.: "Alle Wege führen nach Rom: ein Vergleich der regulatorischen Anforderungen und Medizinprodukte in Europa und den USA", in: Medizin Produkte Journal 18( 2011), 4, S. 243–254.
Schulz, C.: "Corporate Finance für den Mittelstand", in: J. HILSE; W. NETZEL; D. SIMMERT (Ed.): Praxishandbuch Firmenkundengeschäft: Geschäftsfelder Risikomanagement Marketing, Wiesbaden u.a.: Gabler 2010, S. 97–107.
```

That is the reference list entries should have the author list in bold, journal articles should have their title quoted using German-style quotation marks, and editor lists should be given in small caps. The first answerer to the OP's question gives an example database file

```
@BOOK { ko,
  title = {Wissenschaftlich schreiben leicht gemacht},
  publisher = {Haupt},
   year = \{2011\},\
  author = {Kornmeier, M.},
  number = \{3154\},
  series = {UTB},
  address = {Bern},
  edition = \{4\},
   subtitle = {für Bachelor, Master und Dissertation}
@BOOK { nial,
  title = {Wissenschaftliche Arbeiten schreiben mit Word 2010},
   publisher = {Addison Wesley},
  year = {2011},
  author = {Nicol, N. and Albrecht, R.},
  address = {München},
   edition = \{7\}
@ARTICLE { shome,
  author = {Scholz, S. and Menzl, S.},
  title = {Alle Wege führen nach Rom},
   journal = {Medizin Produkte Journal},
   year = \{2011\},
  volume = \{18\},
   pages = \{243-254\},
  subtitle = {ein Vergleich der regulatorischen Anforderungen und Medizinprodukte
  in Europa und den USA},
  issue = {4}
@INBOOK { shu,
   author = {Schulz, C.},
   title = {Corporate Finance für den Mittelstand},
  booktitle = {Praxishandbuch Firmenkundengeschäft},
   year = {2010},
   editor = {Hilse, J. and Netzel, W and Simmert, D.B.},
  booksubtitle = {Geschäftsfelder Risikomanagement Marketing},
  publisher = {Gabler},
  pages = \{97-107\},
   location = {Wiesbaden}
```

that makes use of separate formatting instructions for titles and subtitles, so that an appropriate style template is

(continues on next page)

6.5. Example 5 49

```
SPECIAL-TEMPLATES:
citelabel = <citenum>
sortkey = <citenum>
```

producing the formatted result

Literatur

- 1 Nicol, N.; Albrecht, R.:, Wissenschaftliche Arbeiten schreiben mit Word 2010, 7 Aufl., München u.a.: Addison Wesley 2011.
- 2 Kornmeier, M.:, Wissenschaftlich schreiben leicht gemacht, 4 Aufl., (UTB; 3154), Bern u.a.: Haupt 2011.
- 3 Scholz, S.; Menzl, S.: "Alle Wege führen nach Rom", in: Medizin Produkte Journal 18 (2011), 4 S. 243–254 (2011).
- 5 Schulz, C.: "Corporate Finance für den Mittelstand", in: HILSE, J.; NET-ZEL, W.; SIMMERT, D. (Ed.): Praxishandbuch Firmenkundengeschäft: Geschäftsfelder Risikomanagement Marketing, Wiesbaden u.a.: Gabler 2010, S. 97–107.

6.6 Example 6

The following example is taken from a question posted at http://tex.stackexchange.com/questions/147932/peerage-titles-in-the-author-field-in-bibtex. The question's answerer responds with a Biblatex solution, in which nameaddon and shortauthor fields are added to the *.bib database file, as in

```
@article{Rayleigh1892,
    author = {Strutt, John William},
    journal = {Philosophical Magazine},
    nameaddon = {3rd Baron Rayleigh},
    pages = {481-502},
    shortauthor = {Rayleigh},
    title = {On the influence of obstacles arranged in rectangular order upon the_
properties of a medium},
    volume = {34},
    year = {1892}
}
```

The formatting template makes use of these additional fields, and provides a citation label using the shortauthor field's first three characters followed by the last two characters in the year field:

The formatted result looks like

References

[Ray92] John William Strutt (3rd Baron Rayleigh), "On the influence of obstacles arranged in rectangular order upon the properties of a medium". In: Philosophical Magazine 34 (1892), pp. 481–502.

6.7 Example 7

The following example is taken from a question posted at http://tex.stackexchange.com/questions/145038/some-citation-numbers-in-bold-others-not/168233#168233, where the OP asks if there is a way to have some citations given in **bold** font. This can be achieved with a user-defined script, as shown below. With the database file

```
@Article{refA,
 title = {Can quantum-mechanical description of physical reality be considered,
→complete?},
 author = {A. Einstein and B. Podolsky and N. Rosen},
  journal = {Phys.\ Rev.},
 volume = {47},
 year = \{1935\},
  pages = \{777 - -780\}
@Book{refB,
  title = {Probability Theory: The Logic of Science},
 year = \{2003\},
  author = {E. T. Jaynes},
  publisher = {Cambridge University Press}
@Incollection{refC,
 title = {On the present status of the radiation problem},
  author = {Albert Einstein},
  booktitle = {The Collected Papers of Albert Einstein},
  publisher = {Princeton University Press},
  year = \{1989\},
  pages = \{357 - 375\},
  volume = \{2\}
@Mastersthesis{refD,
 title = {Optical aberration coefficients},
  author = {Matthew Peter Rimmer},
  school = {University of Rochester},
  year = \{1963\}
@Phdthesis{refBOLD,
  title = {On the theory of quanta},
  author = {Louis-Victor de Broglie},
  school = {University of Paris},
  year = \{1925\}
```

and the style template file

6.7. Example 7 51

```
TEMPLATES:
article = <au>, \enquote{<title>,} \textit{<journal>} \textbf{<volume>},{ }...
         [<startpage>--<endpage>|<startpage>|<eid>|] (<year>).
book = <au>, \textit{<title>} (<publisher>, <year>).
incollection = <au>, \enquote{<title>,} in \textit{<booktitle>}, [Chap.~<chapter>, ]..
        [pp.~<startpage>--<endpage>|p.~<startpage>|<eid>|] (<publisher>, <year>).
mastersthesis = <au>, \enquote{<title>,} M.S. dissertation (<school>, <year>).
phdthesis = <au>, \enquote{<title>,} Ph.D. dissertation (<school>, <year>).
SPECIAL-TEMPLATES:
authorlist = <author.to_namelist()>
editorlist = <editor.to_namelist()>
au = <authorlist.format_authorlist()>
ed = <editorlist.format_editorlist()>
OPTIONS:
                           ## whether to allow user scripts in BST files
allow_scripts = True
VARIABLES:
citelabel = create_citelabel(entry, options)
DEFINITIONS:
def create_citelabel(entry, options):
   numnames = len(entry['authorlist'])
   if (numnames == 0): return(entry['citenum'])
    foundit = False
   for name in entry['authorlist']:
        if (name['last'] == 'Broglie') and (name['prefix'] == 'de'):
            foundit = True
           break
   if foundit:
        return(r'\textbf{' + entry['citenum'] + '}')
    else:
        return (entry['citenum'])
```

we can make use of a simple main .tex file

```
\documentclass{article}
\usepackage{cite}
%\makeatletter
% \renewcommand{\@biblabel}[1]{#1.}
%\makeatother
\begin{document}

Some text where I cite papers, where I want to highlight all publications by de_
\Broglie~\cite{refA, refB, refC, refBOLD, refD}.

It would be nice for the citations to to appear sorted like [1-3,{\textbf 4},5].

\bibliography{example7}
\bibliographystyle{example7}
\end{document}
```

to produce a formatted result

Some text where I cite papers, where I want to highlight all publications by de Broglie [4,1-3,5]. It would be nice for the citations to to appear sorted like [1-3,4,5].

References

- A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.* 47, 777–780 (1935).
- [2] E. T. Jaynes, Probability Theory: The Logic of Science (Cambridge University Press, 2003).
- [3] A. Einstein, "On the present status of the radiation problem," in *The Collected Papers of Albert Einstein*, pp. 357–375 (Princeton University Press, 1989).
- [4] L.-V. de Broglie, "On the theory of quanta," Ph.D. dissertation (University of Paris, 1925).
- [5] M. P. Rimmer, "Optical aberration coefficients," M.S. dissertation (University of Rochester, 1963).

6.8 Example 8

The following example is taken from a question posted at http://tex.stackexchange.com/questions/169300/modify-plain-bst-file/172828#172828, where the OP asks for a way to customize the formatted reference list so that the first line contains the author names and date, while the second and subsequent lines are given a hanging indent. This can be achieved with the following setup. With the database file

```
@article { Abrahams1986,
  title = {Scaling description of the dielectric function near the mobility edge},
  author = {Abrahams, Elihu and Lee, P. A.},
  journal = {Phys. Rev. B},
  volume = {33},
  issue = \{2\},
  pages = \{683 - -689\},
  year = \{1986\},
  month = \{1\},
  publisher = {American Physical Society},
  url = {http://link.aps.org/doi/10.1103/PhysRevB.33.683}
@book { Nakayama2003,
  title = {Fractal Concepts in Condensed Matter Physics},
  series = {Springer Series in Solid-State Sciences},
  volume = {140},
  author = {Nakayama, Tsuneyoshi and Yakubo, Kousuke},
  address = {New York},
  publisher = {Springer},
  issn = \{0171-1873\},
  year = \{2003\}
@article { Chalker1988,
  title = {Scaling, diffusion, and the integer quantized Hall effect},
  author = {Chalker, J. T. and Daniell, G. J.},
  journal = {Phys. Rev. Lett.},
```

(continues on next page)

6.8. Example 8 53

```
volume = \{61\},
  issue = \{5\},
  pages = \{593 - 596\},
  year = \{1988\},
  month = \{8\},
  publisher = {American Physical Society},
  url = {http://link.aps.org/doi/10.1103/PhysRevLett.61.593}
@article {Chalker1990,
  journal = {Physica A: Statistical Mechanics and its Applications},
  volume = {167},
  issue = \{1\},
 month = \{8\},
  year = \{1990\},
  day = \{1\},\
  pages = \{253-258\},
  title = {Scaling and eigenfunction correlations near a mobility edge},
  author = {J. T. Chalker}
@article { Pook 1991,
 year = \{1991\},
  issn = \{0722 - 3277\},
  journal = {Zeitschrift für Physik B Condensed Matter},
  volume = \{82\},
  number = \{2\},
  title = {Multifractality and scaling in disordered mesoscopic systems},
  url = {http://dx.doi.org/10.1007/BF01324339},
  publisher = {Springer-Verlag},
  author = {Pook, Werner and Janßen, Martin},
  pages = \{295-298\}
```

and the style template file

```
TEMPLATES:
article = \text{textbf}\{\text{au}, (\text{year}[-\text{month.monthabbrev}()]:}\ [\href{\text{url}}{}
→<title>}|\href{<doi>}{<title>}|<title>}|, \textit{<journal>} \textit{<volume>}([
→ <number> | <issue>|]), [<startpage>--<endpage> | <startpage> | <eid>|].
book = \text{$$\text{book} = \text{$$\text{cau}(sau) = d}$}, (syears[-smonth.monthabbrev()s[-sdays]]):}\\ \\ \text{$$\text{textif}(sau) = \text{$$\text{cau}(sau) = d}$}.
→<title>}[, <edition_ordinal>~ed.][, <series>][ <issn>], <address>: <publisher>.
SPECIAL-TEMPLATES:
authorlist = <author.to_namelist()>
editorlist = <editor.to_namelist()>
authorname.n = [<authorlist.n.prefix> ] <authorlist.n.last>, <authorlist.n.first.</pre>
→initial()>.[ <authorlist.n.middle.initial()>.][, <authorlist.n.suffix>.]
au = \langleauthorname.0\rangle, ...{ \langle \& \rangle \langleauthorname.9\rangle
editorname.n = [<authorlist.n.prefix> ]<authorlist.n.last>, <authorlist.n.first.
→initial()>.[ <authorlist.n.middle.initial()>.][, <authorlist.n.suffix>.]
ed = <editorname.0>, ...{ \& }<editorname.5>
citelabel = <authorlist.0.last>-<year>
sortkey = <citelabel>
```

we can make use of a simple main .tex file

```
\documentclass{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[paper=letterpaper, text={5.8in,8.5in},centering]{geometry}
\usepackage[colorlinks=True,urlcolor=blue,citecolor=black,breaklinks=true]{hyperref}
\usepackage{enumitem}
\makeatletter
   \renewcommand{\@biblabel}[1]{}
   \renewenvironment{thebibliography}[1]
      {\section*{\refname}%
         \@mkboth{\MakeUppercase\refname}{\MakeUppercase\refname}%
         \begin{enumerate}[label={},itemindent=*,leftmargin=3em]
         \@openbib@code
         \sloppy
         \clubpenalty4000
         \@clubpenalty \clubpenalty
         \widowpenalty4000
         \sfcode`\.\@m}
         {\def\@noitemerr
            {\@latex@warning{Empty `thebibliography' environment}}%
\makeatother
\begin{document}
{\noindent}Citations: \cite{Abrahams1986,Nakayama2003,Chalker1988,Chalker1990,
\rightarrowPook1991}
\bibliography{example8}
\bibliographystyle{example8}
\end{document}
```

to produce a formatted result

Citations: [Abrahams-1986, Nakayama-2003, Chalker-1988, Chalker-1990, Pook-1991]

References

```
Abrahams, E. & Lee, P. A., (1986-Jan):
Scaling description of the dielectric function near the mobility edge, Phys. Rev. B 33(2), 683-689.
Chalker, J. T. & Daniell, G. J., (1988-Aug):
Scaling, diffusion, and the integer quantized Hall effect, Phys. Rev. Lett. 61(5), 593-596.
Chalker, J. T., (1990-Aug-1):
Scaling and eigenfunction correlations near a mobility edge, Physica A: Statistical Mechanics and its Applications 167(1), 253-258.
Nakayama, T. & Yakubo, K., (2003):
Fractal Concepts in Condensed Matter Physics, Springer Series in Solid-State Sciences 0171-1873, New York: Springer.
Pook, W. & Janßen, M., (1991):
Multifractality and scaling in disordered mesoscopic systems, Zeitschrift für Physik B Condensed Matter 82(2), 295-298.
```

6.9 Example 9

The following example is taken from a question posted at http://tex.stackexchange.com/questions/172444/suppress-month-and-day-for-journal-articles-and-maintain-for-a-newspaper-article,

6.9. Example 9 55

where the OP asks for a way to create a reference list in which newspaper articles have the month and day of publication shown while journal articles do not. The cleanest way of implementing this in Bibulous is simply to use a different entrytype for newspaper articles and journal articles, as shown below. For this example, we have a database file

```
@newsarticle { test1,
 title = {Ideas \& trends; politics of the web: {Meet}, greet, segregate, meet again}
 url = {http://www.nytimes.com/2004/01/25/weekinreview/ideas-trends-politics-of-the-
→web-meet-greet-segregate-meet-again.html},
 urldate = \{2014-01-14\},\
  journal = {New York Times},
  author = {Harmon, Amy},
  day = \{1\},
 month = jan,
  year = \{2004\}
@article {test2,
  title = {Social movement networks virtual and real},
  volume = \{3\},
  issn = \{1369-118X\},
  url = {http://www.tandfonline.com/doi/abs/10.1080/13691180051033333},
  doi = \{10.1080/13691180051033333\},
  number = \{3\},
  urldate = \{2013-05-31\},
  journal = {Information, Communication \& Society},
  author = {Diani, Mario},
 month = jan,
 year = \{2000\},\
  pages = \{386 - 401\}
```

and a style template file

that when compiled by the main .tex file

```
\documentclass{article}
\usepackage[paper=letterpaper, text={6.2in,8.5in},centering]{geometry}
\usepackage[colorlinks=True,urlcolor=blue,citecolor=black,breaklinks=true]{hyperref}
```

```
\makeatletter
  \renewcommand{\@biblabel}[1]{}
  \renewcommand{\@cite}[2]{{#1\if@tempswa , #2\fi}}
\makeatother

\begin{document}

Lorem \cite{test1} ipsum \cite{test2}.

\bibliography{example9}
\bibliographystyle{example9}

\end{document}
```

produces the formatted result

Lorem (Harmon, 2004) ipsum (Diani, 2000).

References

Harmon, A. (2004, January 1). Ideas & trends; politics of the web: Meet, greet, segregate, meet again. New York Times. Retrieved 2014-01-14, from http://www.nytimes.com/2004/01/25/weekinreview/ideas-trends-politics-of-the-web-meet-greet-segregate-meet-again.html.

Diani, M. (2000). Social movement networks virtual and real. Information, Communication & Society 3(3), 386–401. doi:10.1080/13691180051033333

6.10 Example 10

The following example is taken from a question posted at http://tex.stackexchange.com/questions/68080/beamer-bibliography-icon?lq=1, where the OP asks for a way to place images alongside each bibliography entry. For this example, we have a database file

```
@Book { Lamport-1994,
  author = {Leslie Lamport},
  title = {LaTeX: A Document Preparation System: User's Guide and Reference Manual},
  edition = \{2\},
  year = \{1994\},
  image_filename = {example10_lamport},
  publisher = {Pearson / Prentice Hall}
@Book { Knuth-1984,
 title = {TeXbook},
 year = \{1984\},
 author = {Donald Knuth},
  image_filename = {example10_knuth},
  publisher = {Addison-Wesley}
@Book { Gratzer-2007,
  title = {More Math into LaTeX},
  year = \{2007\},
  author = {George Gr{\"a}tzer},
```

(continues on next page)

6.10. Example 10 57

```
edition = {4},
image_filename = {example10_gratzer},
publisher = {Springer}
}
```

and a style template file

that when compiled by the main .tex file

```
\documentclass{article}
\usepackage[pdftex]{graphicx}
\usepackage[paper=letterpaper, text={4in,5in},centering]{geometry}

\makeatletter
  \renewcommand{\@biblabel}[1]{}
  \newcommand{\citenum}[1]{\def\@cite##1##2{##1}\cite{#1}}

\makeatother

\begin{document}

References \citenum{Knuth-1984}, \citenum{Lamport-1994}, and \citenum{Gratzer-2007}_\citenum \text{-are useful for learning {\TeX} and \LaTeX.}

\bibliography{example10}
\bibliographystyle{example10}
\end{document}

\end{document}
```

produces the formatted result

References 1, 2, and 3 are useful for learning TEX and LATEX.

References



D. Knuth, TeXbook. Addison-Wesley, 1984.



L. Lamport, LaTeX: A Document Preparation System: User's Guide and Reference Manual, 2nd Edition. Pearson / Prentice Hall, 1994.



G. Grätzer, More Math into LaTeX, 4th Edition. Springer, 2007.

6.11 Example 11

The following example shows how to make use of natbib for inline citation commands such as citep and citet. This example uses the database file

```
@article{dijkstra1959,
  author="Edsger Wybe Dijkstra",
  title="A note on two problems in connexion with graphs",
  journal="Numerische Mathematik",
  year="1959",
  volume="1",
  number="1",
  pages="269-271",
  doi="10.1007/BF01386390",
@article{duque2011b,
  author = {Duque, Juan C. and Anselin, Luc and Rey, Sergio J.},
  title = {The Max-$p$-Regions Problem},
  journal = {Journal of Regional Science},
  volume = {52},
  number = \{3\},
  publisher = {Blackwell Publishing Inc},
  pages = \{397 - -419\},
 year = {2012},
  doi = \{10.1111/j.1467-9787.2011.00743.x\},
```

style template file

```
article = {\sc <au>}, \enquote{<title>,} \textit{<journal>,} [\textbf{<volume>}

→][:<number>] [<startpage>-<endpage>|<startpage>] ([<month>[ <day>], ]<year>).

SPECIAL-TEMPLATES:

(continues on next page)
```

6.11. Example 11 59

and main .tex file

```
\documentclass[11 pt]{amsart}
\usepackage[sort,round]{natbib}
\newcommand{\linklogo}{\includegraphics[height=0.8em]{link.pdf}}
\usepackage[sort,round]{natbib}
\usepackage[sort,round]{na
```

to produce the formatted result

```
citet: Duque, Anselin, et al. (2012)
citep: (Duque, Anselin, et al., 2012)
citet: Dijkstra (1959)
citep: (Dijkstra, 1959)

REFERENCES

DIJKSTRA, E. W., "A note on two problems in connexion with graphs,"
Numerische Mathematik, 1:1 269-271 (1959).

DUQUE, J. C., ANSELIN, L., AND REY, S. J., "The Max-p-Regions Problem," Journal of Regional Science, 52:3 397-419 (2012).
```

CHAPTER

SEVEN

BIBULOUS OVERVIEW

Bibulous is a drop-in replacement for BibTeX that makes use of style templates instead of BibTeX's BST language. The code is written in Python and, like BibTeX itself, is open source.

Bibulous developed out of frustration with the complexity of creating bibliography styles using BibTeX's obscure language, and also from the realization that because bibliographies are highly structured, one should be able to specify them simply and flexibly using a template approach. There should be no need to learn a new language just to build a bibliography style, and specifying a style should take only a matter of minutes.

Bibulous incorporates this template approach, and at the same time implements many of the modern enhancements to BibTeX, such as the ability to work with languages other than English, better support for allowing non-standard bibliography entry types, increased options for author name formatting, and more. And one can use the same basic structures and LaTeX commands to generating each of: a bibliography, a glossary, an annotated bibliography, a list of acronyms and symbols, and more, by specifying a different style template for each case.

Bibulous' "style template" files allow a user to visualize the entire bibliography format structure in a concise way within a single page of text. Moreover, the template is structured with its own Python-like mini-language, intended to allow uses to create flexible formatting instructions quickly and easily. The example below illustrates the simplicity of the format.

The project website, with complete documentation, can be found at http://nzhagen.github.io/bibulous/. The complete sourcecode for the project can be found at https://github.com/nzhagen/bibulous. For any questions, the Bibulous project maintainer can be contacted at and.the.light.shattered@gmail.com.

7.1 Installation

Installing using pip::

pip install bibulous

Instructions for installing Bibulous, and for seamlessly integrating it into your normal LaTeX workflow, are given in the getting started section. Users can also consult the guidelines for writing style templates for further information, and the examples section. A FAQ page is also available.

Another approach is to clone the git repository https://github.com/nzhagen/bibulous and place it into your Python path. For compiling bibliographies, the only file needed within the project is bibulous.py, so that going to https://github.com/nzhagen/bibulous and copying bibulous.py from there is sufficient.

7.2 Example

For a very simple bibliography, consisting of only journal articles and books, a complete style template file may consist of just two lines::

The <variable> notation indicates that the corresponding bibliography entry's field is to be inserted into the template there. The [...|...] notation behaves similar to an if...elseif... statement, checking whether a given field is defined within the bibliography entry. If not defined, then it attempts to implement the instruction in the block following the next | character.

We can read the above article template as indicating the following structure for LaTeX-formatting the cited entry in the bibliography (.bib file). For articles, we first insert the list of author names (formatted according to the default form), followed by a comma. If no author field was found in the bibliography entry, then insert ??? to indicate a missing required field. Next insert a quoted title, followed by an italicized journal name, and a boldface volume number (all of these are required fields). Next, if the pages field was found in the entry's database, then parse the start and end page numbers and insert them here. If the pages field indicates that there was only one page, then use that instead. Or if the pages field is not present, then check to see if the eid is defined, and use it instead. However, if none of these three possibilities are available, then insert the "missing field" indicator, ???. Finally, put the year inside parentheses, and if the note field is defined in the entry, then add that to the end (following the period). If note is not defined, then just ignore it.

One can read the book template similarly, and find that it has different required and optional fields. The simplicity of the format allows one to customize databases to suit any use. For example, to use a bibliography entrytype X instead of book, then all that is necessary is to go into the template file and change book to X. If a user wishes to add a new field, such as translator, then if it has been added to the .bib database file, then all that is needed is to add some text to the template, such as (<translator>, trans.) to insert the field into every bibliography entry that has translator defined for that entrytype.

7.3 Developers

Bibulous is a new project, and so it has until now been a solo effort. Anyone interested in helping out is welcome to join; just send an email to the developers mailing list and we will try to help you get involved and show you the ropes. And, this being the maintainer's first open source project, any suggestions by experienced developers are welcome.

Guidelines for developers are given in developer_guide.rst, and includes an overview of the project's strategy and overall code structure. Note that a bug tracking system has not yet been set up for the project. HTML-based documentation is provided at bibulous/doc/_build/html/index.html, and a corresponding PDF file at bibulous/doc/_build/latex/Bibulous.pdf. The setup.py and MANIFEST.in files provided in the repository base directory are used to create a Python package using the disutils distribution utilities module.

7.4 Contact

For any questions, the current Bibulous project maintainer can be reached at and the light. shattered@gmail.com.

7.5 License

Bibulous is released under the MIT/X11 license, meaning that it is free and open source, and that it can be used without restriction in other programs, commercial or not. The full license is given in the file LICENSE.txt.

7.6 Indices and tables

The link below provides an index to the source code for all functions contained in the Bibulous project.

• genindex

• search

7.6. Indices and tables 63