MIBib\TeX Now Deals with Unicode

Jean-Michel HUFFLEN
DISC — University of Burgundy Franche-Comté
TUG & Bacho\TeX, 29th April 2017
Contents

The New Version

Managing input and output encodings

Changes in calling executable programs

Schedule

Conclusion
MlBib\TeX\ 1.3

Implementation (using Scheme) started in 2004.
MlBib\TeX 1.3

Implementation (using Scheme) started in 2004. Many experiments until 2014.
MlBib\TeX 1.3

Implementation (using Scheme) started in 2004. Many experiment until 2014. Including:

- multilingual features,
- enriched syntax for person names,
- ambitious sort procedures,
- bibliographies for biblatex and Con\TeX t,
- namespaces,
- dealing with uncertain information.
Scheme forever!

A bibliography style is a good example of applying the functional paradigm:
Scheme forever!

A bibliography style is a good example of applying the *functional* paradigm:

\[(\lambda (f2) (f2 \ 1\ 2))\]

... but waiting for a version dealing with Unicode!
Scheme forever!

A bibliography style is a good example of applying the functional paradigm:

\[(\lambda \ (f2) \ (f2 \ 1 \ 2))\]

... but waiting for a version dealing with Unicode!

Some ambitious styles need functions to be programmed in Scheme, but I think there is good synergy within \TeX\ world: programmers can help non-programmers.
The new version (1.4)

Major advantage: can deal with Unicode’s full range.
The new version (1.4)

Major advantage: can deal with Unicode’s full range.

Most encodings are available for input and output.
The new version (1.4)

Major advantage: can deal with Unicode’s full range.

Most encodings are available for input and output.

In fact, *byte-based* encodings, no UTF-16.
The new version (1.4)

Major advantage: can deal with Unicode’s full range.

Most encodings are available for input and output.

In fact, byte-based encodings, no UTF-16.

XML parser and printer $\iff$ syntax usable for bibliography database files.
The new version (1.4)

Major advantage: can deal with Unicode’s full range.

Most encodings are available for input and output.

In fact, *byte-based* encodings, no UTF-16.

*XML* parser and printer $\iff$ syntax usable for bibliography database files.

Saving an article’s bibliography as an *XML* file and re-read it.
Compatibility

All the programs available in the previous version will be included into the present one,
Compatibility

All the programs available in the previous version will be included into the present one, except the program populating HAL open-archives site.
Compatibility

All the programs available in the previous version will be included into the present one,
except the program populating HAL open-archives site.
All the features remain usable.
Directives in .bib files

Examples:

%encoding = utf8
%prefix    = bachotex
...

Directives in .bib files

Examples:

%encoding = utf8
%prefix = bachotex
...

Namespaces ← use them with m1bibtex only!
Directives in `.bib` files

Examples:

```%
%encoding = utf8
%prefix = bachotex
...
```

Namespaces \(\leftarrow\) use them with `mlbibtex` only!

Encodings \(\leftarrow\) I recommend to put them explicitly, even if the program tries to guess them.
Output encoding

... for a generated bibliography.
Output encoding

... for a generated bibliography.

ASCII for \LaTeX{}, unless given by the encoding package
Output encoding

... for a generated bibliography.

ASCII for \LaTeX{}, unless given by the encoding package or option of the \texttt{mlbiblatex} program;

UTF-8 for:

- Con\TeX{},
Output encoding

... for a generated bibliography.

ASCII for \LaTeX{}, unless given by the encoding package or option of the \texttt{mlbiblatex} program;

UTF-8 for:

- Con\TeX{},
- XML file got by the \texttt{mlbib2xml} unless another encoding is given.
Output encoding

... for a generated bibliography.

ASCII for \LaTeX, unless given by the encoding package or option of the mlbiblatex program;

UTF-8 for:

- Con\TeXt,
- XML file got by the mlbib2xml unless another encoding is given.

Easy to program if this is a good idea: fontspec's detection $\iff$ UTF-8.
Interface with Scheme

Files in your home directory, e.g.:

~/.mlbibtex  ~/.mlbibcontext
Interface with Scheme

Files in your home directory, e.g.:

```
~/.mlbibtex  ~/.mlbibcontext
```

Change default conventions, e.g.:

```
((encodings-pv 'set-default-4-bib-files)
 'utf8)
```
Allowing XML files to be read, that is:

\[ f \rightarrow f.bib \text{ (kpathsea)}, f-mlbiblio.xml \]
Interface with Scheme (con’d)

Allowing XML files to be read, that is:

\[ f \rightarrow f\text{-}\text{bib} (\text{kpathsea}), f\text{-mlbiblio.xml} \]

by:

\[
((\text{bib\text{-}files\text{-}functions\text{-}pv \ 'set})
 (\text{list s\text{-}parse\text{-}bib\text{-}file sxmlh\text{-}get\text{-}mlbiblio\text{-}xml\text{-}file}))
\]
# Executable programs

M\textsc{lib}BIB\textsc{tex}'s kernel $\implies$ mlbibt\textsc{ex} mlbibcontext
mlbibl\textsc{atex} m\textsc{lib}bib2\textsc{x}ml
Executable programs

\textsc{M\textsc{L}B\textsc{i}B\textsc{T}_{\textsc{E}X}}'s kernel \Rightarrow \texttt{mlbibtex} \quad \texttt{mlbibcontext} \\
\texttt{mlbiblatex} \quad \texttt{mlbib2xml}

Additional options:

\texttt{-inexact} \iff \texttt{mlbibtex, mlbib2xml};

\texttt{-publiweb} \iff \texttt{mlbib2xml}. 
Executable programs

\textsc{MlBibTex}'s kernel $\Rightarrow$ mlbibtex mlbibcontext
mlbiblatex mlbib2xml

Additional options:

- \texttt{-inexact} $\Leftarrow$ mlbibtex, mlbib2xml;

- \texttt{-publiweb} $\Leftarrow$ mlbib2xml.

Argument added for mlbiblatex.
Executable programs

\[ \text{MLbibreX's kernel} \quad \rightarrow \quad \text{mlbibtex} \quad \text{mlbibcontext} \]
\[ \text{mlbiblatex} \quad \text{mlbib2xml} \]

Additional options:

- \text{-inexact} \quad \leftrightarrow \quad \text{mlbibtex, mlbib2xml};

- \text{-publiweb} \quad \leftrightarrow \quad \text{mlbib2xml}.

Argument added for mlbiblatex. Argument removed for mlbibcontext.
Present state

Scheme functions ready for end-users,
Present state

Scheme functions ready for end-users,
but I would like to revise the installation procedure,
Present state

Scheme functions ready for end-users,
but I would like to revise the installation procedure,
and the documentation is not fully updated yet.
When?

Summer 2017 ⇐ Marseillan’s version.
The site lifc.univ-fcomte.fr has been closed.
Where?

The site lifc.univ-fcomte.fr has been closed.

CTAN?
Where?

The site lifc.univ-fcomte.fr has been closed.

CTAN? Contact this summer.
Conclusion

The adventure goes on!