

# MIbibt<sub>E</sub>X Now Deals with Unicode

Jean-Michel HUFFLEN

DISC — University of Burgundy Franche-Comté  
TUG & Bacho<sub>E</sub>X, 29th April 2017

# Contents

The New Version

Managing input and output encodings

Changes in calling executable programs

Schedule

Conclusion

# MLBibT<sub>E</sub>X 1.3

Implementation (using Scheme) started in 2004.

# MLBibT<sub>E</sub>X 1.3

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

# MLBibT<sub>E</sub>X 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 ConT<sub>E</sub>Xt,
- 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:

```
(λ (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 T<sub>E</sub>X 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  $\Leftarrow$  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  $\Leftarrow$  use them with `mlbibtex` 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 `mlbiblatex` program;

UTF-8 for:

- ConT $\text{E}$ Xt,

# 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 $\TeX$ t,
- XML file got by the `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 $\TeX$ t,
- XML file got by the `mlbib2xml` unless another encoding is given.

Easy to program if this is a good idea: `fontspec`'s detection  $\implies$  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)
```

# Interface with Scheme (con'd)

Allowing XML files to be read, that is:

$$f \longrightarrow f.bib \text{ (kpathsea), } f\text{-mlbiblio.xml}$$

# Interface with Scheme (con'd)

Allowing XML files to be read, that is:

$$f \longrightarrow f.bib \text{ (kpathsea)}, f\text{-mlbiblio.xml}$$

by:

```
((bib-files-functions-pv 'set)
 (list s-parse-bib-file sxmlh-get-mlbiblio-xml-file))
```



# Executable programs

MIBIB $\text{\TeX}$ 's kernel  $\implies$       mlbibtex      mlbibcontext  
                                 mlbiblatex      mlbib2xml

Additional options:

-inexact  $\iff$  mlbibtex, mlbib2xml;

-publiweb  $\iff$  mlbib2xml.

# Executable programs

MIBIB $\text{\TeX}$ 's kernel  $\implies$     mlbibtex        mlbibcontext  
                                 mlbiblatex    mlbib2xml

Additional options:

-inexact  $\Leftarrow$  mlbibtex, mlbib2xml;

-publiweb  $\Leftarrow$  mlbib2xml.

Argument added for mlbiblatex.

# Executable programs

MIBIB $\text{\TeX}$ 's kernel  $\implies$     mlbibtex            mlbibcontext  
                                 mlbiblatex       mlbib2xml

Additional options:

-inexact  $\iff$  mlbibtex, mlbib2xml;

-publiweb  $\iff$  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  $\Leftarrow$  Marseillan's version.

# Where?

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!