

The bidi package

François Charette

2007/04/30 v0.2.2

Contents

1 Basic direction switching commands	1	left	4
1.1 Direction of plain footnotes	2	5 Implementation	5
2 Colored text	2	5.1 Basic bidi macros	5
3 Alphabetic numerals in Arabic, Hebrew and Syriac	3	5.2 Logos	6
3.1 Examples	3	5.3 Redefinition of internal L ^A T _E X macros	7
4 Scripts written from right-to-		5.4 Footnotes	9
		5.5 Bidi switching commands	9
		5.6 Footnotes (cont.)	9
		5.7 \textcolor{.}{}	10
		5.8 Alphanumerical numbers	11

1 Basic direction switching commands

The bidi package provides a useful interface for typesetting bidirectional texts with X_EL^AT_EX.¹ Most of its macros are directly copied or adapted from the Babel bidirectional module for Hebrew.²

\RL For short right-to-left insertions in left-to-right environments, \RL{...} \LR does the job, as does \LR{...} in the opposite case. To typeset a full paragraph

¹Actually, bidi can also be used with any ε-T_EX-based engine, notably pdfL^AT_EX, but at this time I do not see how it could be useful outside of the X_EL^AT_EX world.

²See section 60.4 of the *Babel user documentation*.

```
\setRL right-to-left, use the switch \setRL, and to revert to left-to-right, use \setLR  
\setLR (or equivalently \unsetRL).
```

```
\setRL  
This is a right-to-left paragraph.  
\setLR
```

³paragraph. right-to-left a is This

1.1 Direction of plain footnotes

`\setfootnoteRL` With `\setfootnoteRL`, one can insert RL footnotes in LR mode,⁴ and conversely with `\setfootnoteLR` (or equivalently `\unsetfootnoteRL`).

Important note: Currently, X_ET_EX's (version 0.996) placement of the footnote mark at the beginning of plain footnotes in RL context is incorrect (as you can see yourself below).⁵ This will likely be fixed in the next version.

With the package `manyfoot`, it is possible to have several layers of footnotes, each of which can be set in a different direction. See the documentation to ArabX_ET_EX for more details. This should also be possible with `bigfoot`, but I have not tested it yet.

Some other L^AT_EX macros are redefined internally to ensure their correct behaviour in bidirectional display. Contrary to the Babel approach, we leave the commands defined in standard L^AT_EX classes untouched, as we think this should be done by class authors themselves (who are of course encouraged to make them bidi-compatible :-).

2 Colored text

Because of the way bidirectionality is implemented in ε-T_EX, no information on directionality is included in the dvi (or xdv) output. Consequently, if a color `\special` is inserted in the middle of a paragraph, the color change will

³Note that when processing fonts that are not TFM-based, X_ET_EX, contrary to T_EX, uses (roughly speaking) one box per word, instead of one per character. The command `\RL` is thus usually superfluous when inserting a single Arabic or Hebrew word.

⁴ ملاحظات من اليمين إلى اليسار

⁵More precisely, this is a bug in X_ET_EX's handling of glue in right-to-left context.

be effected as though the text would run from left to right. For `\textcolor`, the output will be correct only if the textual argument falls on a single line. A better approach is to set the color of text by means of X_ET_EX's font features, as in this example:

3 Alphabetic numerals in Arabic, Hebrew and Syriac

For convenience the following macros are provided for generating numbers in alphanumerical notation in Arabic (in both Eastern and Western styles) and Hebrew. The interested reader can find information on these systems in the following two Wikipedia articles: http://en.wikipedia.org/wiki/Abjad_numerals and http://en.wikipedia.org/wiki/Hebrew_numerals. Note that the Hebrew implementation in bidi is less sophisticated than the one in Babel, where various special cases are taken into account.

3.1 Examples

- | | |
|----------------|---|
| \abjad | The Arabic abjad numbers can be generated with the command \abjad. Example: \abjad{1863} yields . |
| | In the Maghrib the conventions are somewhat different, and the maghribi forms of the abjad numerals are obtained with the \abjadmaghribi command. |
| \abjadmaghribi | Example: \abjadmaghribi{1863} yields . |
| \hebrewnumeral | Hebrew numerals are generated with the command \hebrewnumeral. Example: \hebrewnumeral{1863} yields . |
| \abjadsyriac | Experimental support is also provided for Syriac abjad numerals, which can be generated with \abjadsyriac. Example: \abjadsyriac{463} yields . |

Note that the above macros only outputs the Unicode characters: the active font must thus contain the necessary glyphs.

4 Scripts written from right-to-left

The Arabic and Hebrew scripts are of course not the only ones to be written from right-to-left. The following is a list of historical⁶ and contemporary scripts that are written primarily or exclusively from right-to-left.⁷ Those still in use today are marked with an asterisk.

Aramaic and Sogdian	Nabatean
Arabic*	N'Ko*
Avestan	Old Hungarian
Berber (ancient)	Orkhon (Turkic runes)
Cypriot Syllabary	Pahlavi and Parthian
Etruscan and Oscan	Phoenician
Hebrew*	Proto-Sinaitic
Iberian (Northern)	Samaritan
Kharosthi	South Arabian and Thamudic
Lydian	Syriac*
Mandaean	Thâna*
Meroitic (cursive)	

Some examples (random glyphs for Kharosthi and Cypriot):

- Old Italic (Etruscan):⁸
- N'ko:⁹

MIṂṂṂṂ ḤṂṂṂ
-IΔΩ□ YΠ Y

⁶The historical right-to-left scripts that are part of Unicode 5.0 are: Cypriot, Old Italic (i.e., Etruscan), Kharosthi and Phoenician.

⁷See <http://www.omniglot.com/writing/direction.htm#rtl>, <http://www.unicode.org/iso15924/>, and <http://www.ancientscripts.com/ws.html>.

⁸The Unicode standard 5.0 specifies left-to-right directionality for the Old Italic script, contrary to historical evidence but in accordance to conventional scholarly usage. To display from right to left the Unicode RIGHT-TO-LEFT OVERRIDE character U+202E must be inserted before each word. This could be easily automated by means of a TECKit font mapping. Etruscanists interested in this can contact me directly.

⁹As can be seen in this example, the ICU rendering engine does not yet support contextual substitutions for N'Ko.

· Thāna:	ଥାନା
· Phoenician:	ଫୋନୀଜିଆ
· Kharosthi:	ଖରସ୍ତି
· Cypriot:	କ୍ୟପ୍ରିଓଟିଆ

5 Implementation

```

1 \NeedsTeXFormat{LaTeX2e}
2 \ProvidesPackage{bidi}
3 [2007/04/30 v0.2.2 Bidirectional typesetting in XeLaTeX]
4 % this is not used for now:
5 \%newif\if@rlmain
6 \%@\rlmainfalse
7 \%DeclareOption{rldocument}{\@rlmaintrue}
8 \%ExecuteOptions{rldocument}
9 \%ProcessOptions
10 %%%%%%%%%%%%%%
11 \ifx\TeXeTstate\undefined\else%
12   \TeXeTstate=1
13 \fi
14 \ifx\beginL\@undefined%
15   \newlinechar`\^J
16   \typeout{\^JTo avoid this error message,\^J%
17     run TeX--XeT or e-TeX engine instead of regular TeX.\^J}
18   \errmessage{Right-to-Left Support Error: use TeX--XeT or e-TeX
19   engine}%
20 \fi
21 \newif\if@rl
22 %\AtBeginDocument{\% Here we set the main document direction
23 % \if@rl\@rlmaintrue%
24 % \else@\rlmainfalse%
25 % \fi%
26 %}

```

5.1 Basic bidi macros

This defines the macros for RL insertion in LR context and conversely. Most of it is taken from the file `rlbabel.def` in the Babel distribution, with minor

changes.

```
27 \let\n@xt=\
28 \def\LR{\protect\pLR}%
29 \def\pLR{\protect\afterassignment\moreLR \let\n@xt= }
30 \def\moreLR{\brace{text \aftergroup\endL \beginL@rlfalse}
31 \def\RL{\protect\pRL}
32 \def\pRL{\protect\afterassignment\moreRL \let\n@xt= }
33 \def\moreRL{\brace{text \aftergroup\endR \beginR@rltrue}
34 \def\brace{text{\ifcat\n@xt{\else\ifcat\n@xt}\fi
35 \errmessage{Missing left brace has been substituted}\fi \bgroup}
36 \everydisplay{\if@rl\aftergroup\beginR\fi }
```

For global right-to-left typesetting, the directional primitive must be included within an `\everypar` construct. The traditional approach was to redefine `\everypar` as a one-argument macro, but since some packages (such as `ledmac`) also play tricks with `\everypar`, this can become problematic. Hence we have adopted Youssef Jabri's clever approach in his `Arabi` package.

```
37 \def\rl@everypar{\if@rl{\setbox\z@\lastbox\beginR\usebox\z@\}\fi}
38 \let\o@everypar=\everypar
39 % The following, surely, will not play well with all packages, e.g. led-
   mac
40 %\def\everypar#1{\o@everypar{\rl@everypar#1}}
41 % But we now use this instead, taken from Arabicore.sty in the Arabi pack-
   age.
42 % Thanks to Youssef Jabri! Shukran jazilan!
43 \newtoks\n@everypar
44 \n@everypar\expandafter{\the\o@everypar}
45 \o@everypar{\rl@everypar\the\n@everypar}
46 \let\everypar=\n@everypar
```

This is not used at the moment:

```
47 \def\@ensure@RL#1{\if@rl#1\else\RL{\#1}\fi}
48 \def\@ensure@LR#1{\if@rl\LR{\#1}\else#1\fi}
```

5.2 Logos

The following redefinitions ensure proper typesetting of the various logos.

```
49 %preserve logos
```

```

50 \let\@@TeX\TeX
51 \def\TeX{\LR{\@@TeX}}
52 \let\@@LaTeX\LaTeX
53 \def\LaTeX{\LR{\@@LaTeX}}
54 \let\@@LaTeXe\LaTeXe
55 \def\LaTeXe{\LR{\@@LaTeXe}}
56 \@ifpackageloaded{xltextra}{
57 \let\@@XeTeX\XeTeX
58 \def\XeTeX{\LR{\@@XeTeX}}
59 }{}
```

5.3 Redefinition of internal L^AT_EX macros

```

60 %VARIOUS LATEX MACROS
61 \def\list#1#2{%
62   \ifnum \@listdepth >5\relax
63     \atodeep
64   \else
65     \global\advance\@listdepth\@ne
66   \fi
67   \rightmargin\z@
68   \listparindent\z@
69   \itemindent\z@
70   \csname @list\romannumeral\the\@listdepth\endcsname
71   \def\@itemlabel{#1}%
72   \let\makelabel\@mklab
73   \nmbrrlistfalse
74   #2\relax
75   \trivlist
76   \parskip\parsep
77   \parindent\listparindent
78   \advance\linewidth -\rightmargin
79   \advance\linewidth -\leftmargin
80   \if@rl
81     \advance\@totallleftmargin \rightmargin
82   \else
83     \advance\@totallleftmargin \leftmargin
84   \fi
```

```

85 \parshape \@ne \@totallleftmargin \linewidth
86 \ignorespaces}
87 %%
88 \def\@dottedtocline#1#2#3#4#5{%
89 \ifnum #1>\c@tocdepth \else
90 \vskip \z@ \oplus .2\p@
91 {\if@rl\rightskip\else\leftskip\fi #2\relax
92 \if@rl\leftskip\else\rightskip\fi \atocrmarg \parfillskip
93 -\if@rl\leftskip\else\rightskip\fi
94 \parindent #2\relax\@afterindenttrue
95 \interlinepenalty\@M
96 \leavevmode
97 \tempdima #3\relax
98 \advance\if@rl\rightskip\else\leftskip\fi \tempdima
99 \null\nobreak\hskip -\if@rl\rightskip\else\leftskip\fi
100 {#4}\nobreak
101 \leaders\hbox{$\m@th
102 \mkern \dotsep mu\hbox{.}\mkern \dotsep
103 mu$}\hfill
104 \nobreak
105 \hb@xt@\pnumwidth{\hfil\normalfont \normalcolor \beginL#5\endL}%
106 \par}%
107 \fi}
108 %%
109 \def\raggedright{%
110 \let\\@centercr
111 \leftskip\z@skip\rightskip\@flushglue
112 \parindent\z@\parfillskip\z@skip}
113 \let\@raggedleft=\raggedleft
114 \let\@raggedright=\raggedright
115 \renewcommand\raggedleft{\if@rl\@raggedright%
116 \else\@raggedleft\fi}
117 \renewcommand\raggedright{\if@rl\@raggedleft%
118 \else\@raggedright\fi}
119 \let\@@underline=\underline
120 \def\underline#1{\@@underline{\if@rl\RL{#1}\else #1\fi}}
121 \if@compatibility
122 \let\undertext=\underline

```

```
123 \fi
```

5.4 Footnotes

This implements the bidirectionality of plain L^AT_EX footnotes.

```
124 \newif\if@rl@footnote
125 \if@rl\@rl@footnotetrue\else\@rl@footnotefalse\fi
126 \let\@@footnoterule=\footnoterule
127 \def\LRfootnoterule{\@@footnoterule}
128 %bug?-->this causes new par to be set, hence hack with vskip :-{ 
129 \def\RLfootnoterule{\vskip -\baselineskip\hb@xt@\hsize{\hss\vbox{\@@footnoterule}}}
130 %
131 \def\setfootnoteRL{\@rl@footnotetrue}
132 \def\unsetfootnoteRL{\@rl@footnotefalse}
133 \def\setfootnoteLR{\unsetfootnoteRL}
134 %
135 \def\footnoterule{\if@rl@footnote\RLfootnoterule\else\LRfootnoterule\fi}
```

5.5 Bidi switching commands

Commands for switching between LR and RL global contexts.

```
136 \def\setRL{\@rltrue\@rl@footnotetrue}
137 \def\unsetRL{\@rlfalse\@rl@footnotefalse}
138 \def\setLR{\unsetRL}
```

5.6 Footnotes (cont.)

This redefines L^AT_EX's internal footnote macros.

```
139 \long\def\@footnotetext#1{\insert\footins{%
140   \if@rl@footnote\@rltrue\else\@rlfalse\fi
141   \reset@font\footnotesize
142   \interlinepenalty\interfootnotelinepenalty
143   \splittopskip\footnotesep
144   \splitmaxdepth \dp\strutbox \floatingpenalty \MM
145   \hsize\columnwidth \parboxrestore
146   \protected@edef\@currentlabel{%
147     \csname p@footnote\endcsname\@thefnmark
148   }%
```

```

149  \color@begingroup
150  \@makefntext{%
151    \rule{z}{\footnotesep}\ignorespaces#1\@finalstrut\strutbox}%
152  \color@endgroup}%
153 %... also in minipages
154 \long\def\@mpfootnotetext#1{%
155  \global\setbox\@mpfootins\vbox{%
156  \if@rl@footnote\@rltrue\else\@rlfalse\fi
157  \unvbox\@mpfootins
158  \reset@font\footnotesize
159  \hsize\columnwidth
160  \parboxrestore
161  \protected@edef\@currentlabel
162    {\csname p@\mpfootnote\endcsname\@thefnmark}%
163  \color@begingroup
164  \@makefntext{%
165    \rule{z}{\footnotesep}\ignorespaces#1\@finalstrut\strutbox}%
166  \color@endgroup
167  }%

```

5.7 \textcolor

\textcolor needs to be redefined so that the color-switching commands in the output will affect the correct portion of text also in RL context, but this only works if the textual argument does not overlap more than one line, which is why we issue a warning. See § 2 above for an alternative approach to colored text using X_ET_EX font features. The following code (less the warning) is again taken from `rlbabel.def`.

```

168 \ifx\@textcolor\@undefined\else%
169 \AtBeginDocument{%
170  \def\@textcolor#1#2#3{%
171  \if@rl\PackageWarning{bidi}{%
172    \textcolor will not work in RL mode if the textual argu-
173    ment spans more than one line.^J%
174    With XeTeX you should set color as a font feature instead.}%
175  \beginL\protect\leavevmode{\color#1{#2}\beginR#3\endR}\endL%
176  \else%
177  \protect\leavevmode{\color#1{#2}#3}%

```

```

177      \fi%
178    }%
179 }%
180 \fi
181 %%%

```

5.8 Alphanumerical numbers

The code that follows is a straightforward adaptation of the corresponding code in the `Arabi` package. The code for Hebrew numerals is a simplified version of the implementation in `Babel`.

```

182 \def\abjad#1{%
183 \ifnum#1>1999 \PackageWarning{bidi}{Illegal value (#1) for abjad nu-
    meral} {#1}
184 \else
185 \ifnum#1<\z@\space\PackageWarning{bidi}{Illegal value (#1) for ab-
    jad numeral}%
186 \else
187 \ifnum#1<10\expandafter\abj@num@i\number#1%
188 \else
189 \ifnum#1<100\expandafter\abj@num@ii\number#1%
190 \else
191 \ifnum#1<\@m\expandafter\abj@num@iii\number#1%
192 \else
193 \ifnum#1<\@M\expandafter\abj@num@iv\number#1%since #1<2000, we must have 1000
194 \fi
195 \fi
196 \fi
197 \fi
198 \fi
199 \fi
200 }
201
202 \def\abjadmaghribi#1{%
203 \ifnum#1>1999 \PackageWarning{bidi}{Illegal value (#1) for abjad nu-
    meral} {#1}
204 \else

```

```

205 \ifnum#1<\z@\space\PackageWarning{bidi}{Illegal value (#1) for ab-
jad numeral}%
206 \else
207 \ifnum#1<10\expandafter\abj@num@i\number#1%
208 \else
209 \ifnum#1<100\expandafter\abj@maghribi@num@ii\number#1%
210 \else
211 \ifnum#1<@\m\expandafter\abj@maghribi@num@iii\number#1%
212 \else
213 \ifnum#1<@\M\expandafter\abj@maghribi@num@iv\number#1%
214 \fi
215 \fi
216 \fi
217 \fi
218 \fi
219 \fi
220 }
221
222 \def\hebrewnumeral#1{%
223 \ifnum#1>9999 \PackageWarning{bidi}{Illegal value (#1) for Hebrew nu-
meral} {#1}%
224 \else
225 \ifnum#1<\z@\space\PackageWarning{bidi}{Illegal value (#1) for He-
brew numeral}%
226 \else
227 \ifnum#1<10\expandafter\heb@num@i\number#1%
228 \else
229 \ifnum#1<100\expandafter\heb@num@ii\number#1%
230 \else
231 \ifnum#1<@\m\expandafter\heb@num@iii\number#1%
232 \else
233 \ifnum#1<@\M\expandafter\heb@num@iv\number#1%
234 \fi
235 \fi
236 \fi
237 \fi
238 \fi
239 \fi

```

```

240 }
241
242 \def\abjad@syriac#1{%
243 \ifnum#1>499 \PackageWarning{bidi}{Illegal value (#1) for Syriac ab-
    jad numeral} {#1}
244 \else
245 \ifnum#1<\z@\space\PackageWarning{bidi}{Illegal value (#1) for Syr-
    iac abjad numeral}%
246 \else
247 \ifnum#1<10\expandafter\abj@syr@num@i\number#1%
248 \else
249 \ifnum#1<100\expandafter\abj@syr@num@ii\number#1%
250 \else
251 \expandafter\abj@syr@num@iii\number#1%
252 \fi
253 \fi
254 \fi
255 \fi
256 }
257 %

```

5.8.1 Symbol for zero

By default the abjad symbol for zero is not defined, but if a package writer or user owns or develops a font containing appropriate glyphs (e.g. as found in medieval manuscripts), it can be easily redefined.

```

258 \def\abjad@zero{}
259 \def\hebnum@zero{}
260 \def\syriac@zero{}

```

5.8.2 Definitions of alphanumeric characters

```

261 \def\abj@num@i#1{%
262 \ifcase#1\or\char"0627 \or\char"0628 \or\char"062C \char"200D\or\char"062F %
263 \or\char"0647 \or\char"0648 \or\char"0632 \or\char"062D \or\char"0637 \fi
264 \ifnum#1=\z@\abjad@zero\fi}
265 \def\abj@num@ii#1{%
266 \ifcase#1\or\char"064A \or\char"0643 \or\char"0644 \or\char"0645 \or\char"0646 %

```

```

267           \or\char"0633 \or\char"0639 \or\char"0641 \or\char"0635 \fi
268 \ifnum#1=\z@\fi\abj@num@i}
269 \def\abj@num@iii#1{%
270 \ifcase#1\or\char"0642 \or\char"0631 \or\char"0634 \or\char"062A \or\char"062B %
271             \or\char"062E \or\char"0630 \or\char"0636 \or\char"0638 \fi
272 \ifnum#1=\z@\fi\abj@num@ii}
273 \def\abj@num@iv#1{%
274 \ifcase#1\or\char"063A \fi
275 \ifnum#1=\z@\fi\abj@num@iii}
276
277 %maghribi و -> و و -> و و -> و و -> ب ب -> خ خ -> و
278 \def\abj@maghribi@num@ii#1{%
279 \ifcase#1\or\char"064A \or\char"0643 \or\char"0644 \or\char"0645 \or\char"0646 %
280             \or\char"0635 \or\char"0639 \or\char"0641 \or\char"0636 \fi
281 \ifnum#1=\z@\abjad@zero\fi\abj@num@i}
282 \def\abj@maghribi@num@iii#1{%
283 \ifcase#1\or\char"0642 \or\char"0631 \or\char"0633 \or\char"062A \or\char"062B %
284             \or\char"062E \or\char"0630 \or\char"0638 \or\char"063A \fi
285 \ifnum#1=\z@\fi\abj@maghribi@num@ii}
286 \def\abj@maghribi@num@iv#1{%
287 \ifcase#1\or\char"0634 \fi
288 \ifnum#1=\z@\fi\abj@maghribi@num@iii}
289
290 %Hebrew alphabetic numerals (primitive implementation -- the code in Ba-
291     bel is more complex):
291 \def\heb@num@i#1{%
292 \ifcase#1\or\char"05D0 \or\char"05D1 \or\char"05D2 \or\char"05D3 \or\char"05D4 %
293             \or\char"05D5 \or\char"05D6 \or\char"05D7 \or\char"05D8 \fi
294 \ifnum#1=\z@\hebnum@zero\fi}
295 \def\heb@num@ii#1{%
296 \ifcase#1\or\char"05D9 \or\char"05DB \or\char"05DC \or\char"05DE \or\char"05E0 %
297             \or\char"05E1 \or\char"05E2 \or\char"05E4 \or\char"05E6 \fi
298 \ifnum#1=\z@\fi\heb@num@i}
299 \def\heb@num@iii#1{%
300 \ifcase#1\or\char"05E7 \or\char"05E8 \or\char"05E9 \or\char"05EA \or\char"05DA %
301             \or\char"05DD \or\char"05DF \or\char"05E3 \or\char"05E5 \fi
302 \ifnum#1=\z@\fi\heb@num@ii}
303 \def\heb@num@iv#1{%

```

```

304 \ifcase#1\or\char"05D0\char"05F3 \or\char"05D1\char"05F3 \or\char"05D2\char"05F3 %
305     \or\char"05D3\char"05F3 \or\char"05D4\char"05F3 \or\char"05D5\char"05F3 %
306     \or\char"05D6\char"05F3 \or\char"05D7\char"05F3 \or\char"05D8\char"05F3 \fi
307 \ifnum#1=\z@\fi\heb@num@iii}
308 %TODO [optionally]
309 % -insert gereshayim (\char"05F4) before last letter of a se-
310 % quence of two letters or more
311 % -insert geresh (\char"05F3) after letter if it is alone
312 \def\abj@syr@num@i#1{%
313 \ifcase#1\or\char"0710 \or\char"0712 \or\char"0713 \or\char"0715 %
314 \or\char"0717 \or\char"0718 \or\char"0719 \or\char"071A \or\char"071B\fi
315 \ifnum#1=\z@\syriac@zero\fi}
316 \def\abj@syr@num@ii#1{%
317 \ifcase#1\or\char"071D \or\char"071F \or\char"0720 \or\char"0721 \or\char"0722 %
318     \or\char"0723 \or\char"0725 \or\char"0726 \or\char"0728 \fi
319 \ifnum#1=\z@\fi\abj@syr@num@i}
320 \def\abj@syr@num@iii#1{%
321 \ifcase#1\or\char"0729 \or\char"072A \or\char"072B \or\char"072C \fi
322 \ifnum#1=\z@\fi\abj@syr@num@ii}

```