a few tips
Table of contents

Introduction 2

1 Setting up fonts 3

2 A mixed layout 6

3 Numbering and positioning 9

4 Tables 10

5 The LUA interface 17

6 Going vertical 18
Introduction

With CONTeXT you can typeset in two directions: from left to right and from right to left. In fact you can also combine these two directions, like this:

There are many \righttoleft \maincolor \it scripts in use} and some run into the other direction. However, there is \righttoleft \maincolor \it no fixed relation \lefttoright \black \it between the} direction of the script} and cars being driven left or right of the road.

There are many esu ni stpircs and some run into the other direction. However, there is dexfi on tpircs eht fo noitcerid between the noitaler and cars being driven left or right of the road.

Even Luigi Scarso pointed out that the fixed reversed into dexif but not in the example where fixed becomes dexfi. This signals an important property of the way the text gets processed: you input something, at some points font features get applied (like ligatures) and in the end the resulting glyph stream is reversed. By that time the combination of f+i has become fi! So, be prepared for surprises.

This manual is written by a left to right user so don’t expect a manual on semitic typesetting. Also don’t expect a (yet) complete manual. I’ll add whatever comes to mind. This is not a manual about Hebrew or Arabic, if only because I can’t read any of those scripts (languages). I leave that to others to cover.

This is work in progress and might always be! So expect errors and typos. As with anything related to typesetting the truth about how it should be done and what looks best is not absolute. So, the most we can offer is flexibility and the way CONTeXT is setup permits that.

Of course this is not possible without input. When we moved to CONTeXT LMTX, the bidi thread was picked up by Mohammad Hossein Bateni, Idris Samawi Hamid, Wolfgang Schuster and myself. So, expect more!

Hans Hagen
Hasselt, NL
So let's see how Arabic and Hebrew come out:

The sentence \textit{I have no clue what this means.} is translated (by Google Translate) into \textit{لاَيَسْ لَدِيَ أيْنِي فَكرَاةً عَمَّا يَعْنِى هَذَا.} which is then translated back to \textit{I have no idea what this means.} so maybe Arabic has no clue what a clue is. The suggested Arabic pronunciation is \textit{lays laday 'ayu fikrat eamaa yaenih hadha}. Hebrew also likes ideas more: \textit{רמואזזהמגשומילןיא}.

The sentence “I have no clue what this means.” is translated (by Google Translate) into “لاَيَس لَدِيَ أيْنِي فَكِرَاةً عَمَّا يَعْنِى هَذَا” which is then translated back to “I have no idea what this means.” so maybe Arabic has no clue what a clue is. The suggested Arabic pronunciation is “lays laday 'ayu fikrat eamaa yaenih hadha”. Hebrew also likes ideas more: “לְמָאֲזָזָה מַגְשּוֹמִילְּןַיא.”

According to Idris Hamid the Arabic should actually be this: “لاَيَس لَدِيَ أيْنِي فَكِرَاةً عَمَّا يَعْنِى هَذَا” and the transliteration “Laysa ladayya ayyu fikratin ‘ammā ya’nihi hādhā”.

The \textsc{con}tex (or any \texttt{TeX}) ecosystem deals with languages and fonts. Languages (that relate to scripts) have specific characteristics, like running from right to left, and fonts provide a repertoire of glyphs and features. There is no real (standard) relationship between these. In for instance browsers, there are automatic fallback systems for missing characters in a font: another font is taken. These fallbacks are often not easy to tweak.

In this document we use Dejavu and although that font has Arabic shapes in its monospace variant, the serifs come without them (at least when I write this down). Before we actually define the bodyfont we hook in some fallbacks. The typescript for Dejavu has lines like this:

\begin{verbatim}
definefontsynonym [SerifBoldItalic] [name:dejavuserifbolditalic] [features=default, fallbacks=SerifBoldItalic]
definefontfallback [Serif] [scheherazaderegular*arabic] [arabic] [force=yes,rscale=1.5] definefontfallback [SerifBold] [scheherazadebold*arabic] [arabic] [force=yes,rscale=1.5] definefontfallback [SerifItalic] [scheherazaderegular*arabic] [arabic] [force=yes,rscale=1.5] definefontfallback [SerifBoldItalic] [scheherazadebold*arabic]
\end{verbatim}

This permits us to do this:

\begin{verbatim}
definefontsynonym [SerifBoldItalic] [name:dejavuserifbolditalic] [features=default, fallbacks=SerifBoldItalic]
definefontfallback [Serif] [scheherazaderegular*arabic] [arabic] [force=yes,rscale=1.5] definefontfallback [SerifBold] [scheherazadebold*arabic] [arabic] [force=yes,rscale=1.5] definefontfallback [SerifItalic] [scheherazaderegular*arabic] [arabic] [force=yes,rscale=1.5] definefontfallback [SerifBoldItalic] [scheherazadebold*arabic]
\end{verbatim}
In addition we set up the languages:

\setuplanguage[ar][font=arabic,bidi=right]
\setuplanguage[he][font=hebrew,bidi=right]

The following example demonstrates what the effects of these commands are:

{اَذٰهِهيِنْعَياَّمَعٍةَرْكِفُّئَاَّيَدَلَسْيَل}.
{راموآ هی هیم نیوم ویتی اآ}.
{\righttoleft اَذٰهِهيِنْعَياَّمَعٍةَرْكِفُّئَاَّيَدَلَسْيَل}.
{\righttoleft راموآ هی هیم نیوم ویتی اآ}.
{\ar \righttoleft اَذٰهِهيِنْعَياَّمَعٍةَرْكِفُّئَاَّيَدَلَسْيَل}.
{\ar \righttoleft راموآ هی هیم نیوم ویتی اآ}.
{\he \righttoleft اَذٰهِهيِنْعَياَّمَعٍةَرْكِفُّئَاَّيَدَلَسْيَل}.
{\he \righttoleft راموآ هی هیم نیوم ویتی اآ}.

لَیِّسَ لَذِیٌّ اَیُّ فِکْرُهُ عَمَّا يَغْنیِهِ هَذَا.
In principle you can also rely on automatic direction changes, for instance by using the following command:

\setupdirections{
bidi=global,  
method=three
}

But that doesn’t do a font switch for you, nor does it do any of the other language related settings. It really helps if you properly tag your document content, as in:

{\arにくいですね。
}

{\heรามואזה枉רשומילןיא}

One reason to set the font parameter for a language is that it will activate the right features in a font. Instead of falling back on some default, we can be very specific in what we want to enable.
2 A mixed layout

The typesetting engine normally works from left to right and top to bottom. Going from right to left actually involved two decisions:

- the direction of the display elements, the paragraphs
- the direction of the inline text, the lines

The first one is kept track of in a state variable. Every paragraph starts with a node that carries, among other information, that state. This node is added automatically and does not interfere with the typesetting. The inline direction is more intrusive as it is marked by nodes that indicate the beginning and end of a reversed strip. This mechanism is rather reliable and normally works out well. Take this:

\left{ \righttoleft right } \left
\left{ \righttoleft right } \left
\left{ \righttoleft right } \left
\left{ \righttoleft right } \left

You can see that we need to be careful with spaces as they can end up inside or outside a substream and by swapping next to each other:

left their left
left their left
left theirleft
left theirleft

We can wrap the lines in boxes as in:

\hbox{left} \space{\bf \righttoleft right} \space{left}
\hbox{left} \space{ \bf space \ righttoleft right} \space{left}
\hbox{left} \space{\bf \righttoleft right space} \space{left}
\hbox{left} \space{ \bf space \ righttoleft right space} \space{left}

When visualize the spaces we get this:

left
left
left
left

The space of a normal and bold font in the same family normally is the same but let’s mix with a larger size:

\hbox{left} \{\bf \righttoleft right \} \left
\hbox{left} \{\bf space \ righttoleft right \} \left
\hbox{left} \{\bf \righttoleft right space\} \left
\hbox{left} \{\bf space \ righttoleft right space\} \left

Now we get the following. As you can see, it really matters where we put the braces.
Once you are accustomed to tagging and \TeX you will probably not fall into these traps. In figure 2.1 we show a large version.

The \texttt{\righttoleft} command actually has two meanings. This can best be seen from an example.

\righttoleft \textbf{How will this come out?}

\righttoleft \textbf{How will this come out?}

And \righttoleft \textbf{how will this come out?}

And \texttt{\righttoleft} \textbf{how will this come out?}

When we start a paragraph (or in \TeX speak: when we are still in vertical mode) the paragraph direction as well as the inline direction is set. Otherwise only the inline direction is set. There are low level \TeX commands (primitives) to set the direction but you can best \textit{not} use these because we need to do a bit more than that.

There are quite some low level commands related to changing directions. Some deal with the layout, some with boxes. We might provide more in the future.
When we talk about layout, we mean the overall layout, concerning the document as a whole. We can have a dominantly l2r, dominantly r2l or mixed setup. In a next chapter we will give more details on the dominant setup. Here we stick to mentioning that the document flow direction is set with

```latex
\setupalign[r2l] % or r2l
```

When a command to setup an environment has a `align` parameter, the same keywords can be uses as part of the specification.\(^1\)

\(^1\) We haven’t tested all situations and possible interferences. Just report anomalies to the mailing list.
3 Numbering and positioning

todo: columns (direction key), numbers (conversion sets), margins (begin/end), etc
4 Tables

Because tables represent a lot of the issues we have to deal with we start with a little introduction. This chapter can therefore also be seen a bit more in depth discussion of the kind of problems that we have to deal with.

It makes no sense to simplify bi-directional typesetting to two variants: left-to-right and right-to-left. Among the reasons are that till now there never was a consistent model presented. Over the years several individual features were introduced but there is no simple truth here. This is quite understandable in the perspective of TeX where all is about flexibility and user control. Here are some reason why we settled on the approach that is now present in CONTeXT LMTX.

A document layout is normally dominated by a main direction. The main language can be (for instance) English so all goes from left to right but snippets can be Arabic or Hebrew. When they are inline, they need to run from left to right for sure. But what about a paragraph, the order of cells in a table, the placement of images at the margin, binding to the edges of the text area in a double sided layout, footnotes? A mixed direction document might want to set up the dominate direction and then specific preferences per structural element.

In CONTeXT we can use the alignment related keywords to deal with the inline and display directions. At the document level that is done with \setupalign and in commands we use the align key(s). When we have a composite structural element the direction key controls the order or flushing the sub-elements. With each combination of these options there can be interesting side effects, hopefully most are predictable and intentional but some might be the result of interaction between features and of course some can be bugs (that then need to be fixed).

Because in tables the cells are basically boxes glued together in rows, it helps to know a bit about what happens at the lowest level. When we enter a snippet of text eventually it will get boxed, either as (part of) a line or as a (more or less) hard coded box, think of something framed, a floating body or a table cell. Inside a box (we’re talking a list then) the stream of glyphs, rules, skips, kerns and whatever else can be mixed with directional directives that push and pop a directional state. At a pretty low level this can be done as follows (keep in mind that CONTeXT has more robust interfaces that work better with other features). First we see a normal l2r box with an embedded r2l snippet.

```
\hbox \bgroup
  abc
\bgroup
  \textdirection\directionrighttoleft
  DEF
\egroup
  ghi
\egroup

abcFEDghi
```

The next example explicitly makes a r2l box that has an opposing snippet:

```
\hbox direction \directionrighttoleft \bgroup
  abc
\bgroup
```
You can imagine that global settings influence this process and the more explicit boxing takes place the more care has to be taken. Often the typesetting completely reverses with a right-to-left switch but just as often we don’t want that. This is why naive switching doesn’t work well: there has to be some integration in the whole system.

We now repeat these two examples but reverse the content. This reversal feature is to be used with care: the node list is literally reversed. This is different from the directional directive because these only pass information to the backend about the way the result is positioned. The reverse keyword can also be given to an alignment command which is quite safe because an alignment row is just a list of packaged cells. Their content is left untouched.

The table related mechanisms in \texttt{CONTeXT} have (l2r and r2l) alignment support in the cells and (reverse) direction support too. We have traditional tables:

```latex
\starttable[[l|l|l]]
<table>
<thead>
<tr>
<th>left</th>
<th>right</th>
<th>left</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>r</td>
<td>l</td>
</tr>
</tbody>
</table>
\stopstable
```

Running text tabulation:

```
\starttabulate[[l|l|l]]
abc labc
```

ihgDEFcba
And natural tables:

\begin{table}
  \begin{tabular}{ll}
    \texttt{\texttt{\texttt{align}=flushleft}} & \texttt{left} \\
    \texttt{\texttt{align}=flushright} & \texttt{right} \\
  \end{tabular}
\end{table}

We've put these definitions in buffers, and call these with:

\begin{itemize}
  \item \texttt{\setup\table\[direction=normal\]} \texttt{\getbuffer\[table:table\]} \texttt{\blank}
  \item \texttt{\setup\table\[direction=reverse\]} \texttt{\getbuffer\[table:table\]} \texttt{\blank}
  \item \texttt{\setup\tabulate\[direction=normal\]} \texttt{\getbuffer\[table:tabulate\]} \texttt{\blank}
  \item \texttt{\setup\tabulate\[direction=reverse\]} \texttt{\getbuffer\[table:tabulate\]} \texttt{\blank}
  \item \texttt{\setup\TABLE\[split=yes\]} \texttt{\getbuffer\[table:TABLE\]} \texttt{\blank}
  \item \texttt{\setup\TABLE\[direction=normal\]} \texttt{\getbuffer\[table:TABLE\]} \texttt{\blank}
  \item \texttt{\setup\TABLE\[direction=reverse\]} \texttt{\getbuffer\[table:TABLE\]} \texttt{\blank}
\end{itemize}

In a lefttoright layout we get:

\begin{itemize}
  \item \texttt{left} \quad \texttt{right} \\
  \quad \texttt{l} \quad \texttt{r} \\
  \texttt{right} \quad \texttt{left} \\
  \quad \texttt{r} \quad \texttt{l} \\
\end{itemize}

In a righttoleft one we get:

\begin{itemize}
  \item \texttt{left} \quad \texttt{right} \\
  \quad \texttt{l} \quad \texttt{r} \\
\end{itemize}
Table environments are kind of complicated because we can have free standing tables and wrapped ones. They can be part of a float and get split if needed. In most cases using a tabulate makes most sense because that one is meant to be part of the text flow. It can handle paragraphs well too.

\starttabulate
| l | p |
\dorecurse {2} {
    \NC sapolsky \NC \samplefile{sapolsky} \NC \NR
}
\stoptabulate

\starttabulate
| l | l | l | l |
\dorecurse {2} {
    \NC sapolsky \NC robert \NC \darkred \tabulateparameter{direction} \NC \NR
}
\stoptabulate

Agriculture is a fairly recent human invention, and in many ways it was one of the great stupid moves of all time. Hunter-gatherers have thousands of wild sources of food to subsist on. Agriculture changed that all, generating an overwhelming reliance on a few dozen domesticated food sources, making you extremely vulnerable to the next famine, the next locust infestation, the next potato blight. Agriculture allowed for stockpiling of surplus resources and thus, inevitably, the unequal stockpiling of them — stratification of society and the invention of classes. Thus, it allowed for the invention of poverty. I think that the punch line of the primate-human difference is that when humans invented poverty, they came up with a way of subjugating the low-ranking like nothing ever seen before in the primate world.

Agriculture is a fairly recent human invention, and in many ways it was one of the great stupid moves of all time. Hunter-gatherers have thousands of wild sources of food to subsist on. Agriculture changed that all, generating an overwhelming reliance on a few dozen domesticated food sources, making you extremely vulnerable to the next famine, the next locust infestation, the next potato blight. Agriculture allowed for stockpiling of surplus resources and thus, inevitably, the unequal stockpiling of them — stratification of society
Agriculture is a fairly recent human invention, and in many ways it was one of the great stupid moves of all time. Hunter-gatherers have thousands of wild sources of food to subsist on. Agriculture changed that all, generating an overwhelming reliance on a few dozen domesticated food sources, making you extremely vulnerable to the next famine, the next locust infestation, the next potato blight. Agriculture allowed for stockpiling of surplus resources and thus, inevitably, the unequal stockpiling of them — stratification of society and the invention of classes. Thus, it allowed for the invention of poverty. I think that the punch line of the primate-human difference is that when humans invented poverty, they came up with a way of subjugating the low-ranking like nothing ever seen before in the primate world.
So called natural tables are also popular. They are coded like HTML tables and formatting is mostly automatic.

\bTABLE
\bTR
  \bTD[align=flushleft]  left  \eTD
  \bTD[align=flushright] right  \eTD
  \bTD[align=flushforward] forward  \eTD
  \bTD[align=flushbackward] backward  \eTD
\eTR
\bTR
  \bTD[align=flushleft]  l  \eTD
<table>
<thead>
<tr>
<th>left</th>
<th>right</th>
<th>forward</th>
<th>backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
<td>r</td>
<td>f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>left</th>
<th>right</th>
<th>forward</th>
<th>backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
<td>r</td>
<td>f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>left</th>
<th>right</th>
<th>forward</th>
<th>backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
<td>r</td>
<td>f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>left</th>
<th>right</th>
<th>forward</th>
<th>backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
<td>r</td>
<td>f</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>left</th>
<th>right</th>
<th>forward</th>
<th>backward</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td></td>
<td>r</td>
<td>f</td>
</tr>
</tbody>
</table>
5 The LUA interface

We assume that you run CONTeXt MkIV in combination with LUA\TeX. Direction support in this engine has been improved over time. Originally the OMEGA (ALEPH) direction model was used but in the meantime it has been stripped to the basics, and what used to be so called whatsits (extension nodes) are now first class nodes. Of the many directions only four are kept in LUA\TeX and they are indicated by three letters:

0 TLT left to right
1 TRT right to left
2 LTL not used in context (obsolete)
3 RTT not used in context (obsolete)

In LUAMETA\TeX, and therefore CONTeXt LMTX we only have the first two. Therefore in LMTX you normally don’t have to worry about checking for them at the LUA end because they are irrelevant for calculations (the vertical ones swapped the horizontal and vertical progression). Also, when really needed, we use the direction keys with numerical indicators, so zero for l2r and one for r2l. These values are used for local par nodes as well as direction nodes. In addition a direction node has a subtype:

0 normal comparable to +
1 cancel comparable to -
6 Going vertical

Normally the term bidi is reserved for horizontal direction swapping but there is no reason to limit our view to that. So, here I will spend some words on how we can deal with vertical directions.

*I will move some (not yet public) explanation from elsewhere to here in due time.*