The LuaTeX-ja package

The LuaTeX-ja project team

20210319.0 (March 19, 2021)
This documentation is far from complete. It may have many grammatical (and contextual) errors. Also, several parts are written in Japanese only.
Part I

User’s manual

1 Introduction

The LuaTEX-ja package is a macro package for typesetting high-quality Japanese documents when using LuaTEX.

1.1 Backgrounds

Traditionally, ASCII pTEX, an extension of TeX, and its derivatives are used to typeset Japanese documents in TeX. pTEX is an engine extension of TeX: so it can produce high-quality Japanese documents without using very complicated macros. But this point is a mixed blessing: pTEX is left behind from other extensions of TeX, especially εTEX and pdfTeX, and from changes about Japanese processing in computers (e.g., the UTF-8 encoding).

Recently extensions of pTEX, namely upTEX (Unicode-implementation of pTEX) and ε-pTEX (merging of pTEX and ε-TEX extension), have developed to fill those gaps to some extent, but gaps still exist.

However, the appearance of LuaTEX-ja changed the whole situation. With using Lua "callbacks", users can customize the internal processing of LuaTEX. So there is no need to modify sources of engines to support Japanese typesetting: to do this, we only have to write Lua scripts for appropriate callbacks.

1.2 Major changes from pTEX

The LuaTEX-ja package is under much influence of pTEX engine. The initial target of development was to implement features of pTEX. However, implementing all feature of pTEX is impossible, since all process of LuaTEX-ja must be implemented only by Lua and TeX macros. Hence LuaTEX-ja is not a just porting of pTEX; unnatural specifications/behaviors of pTEX were not adopted.

The followings are major changes from pTEX. For more detailed information, see Part III or other sections of this manual.

■ Command names   pTEX adds several primitives, such as \kanjiskip, \prebreakpenalty, and \ifydir. They can be used as follows:
\kanjiskip=10pt \dimen0=\kanjiskip
\tbaselineshift=0.1zw
\dimen0=\tbaselineshift
\prebreakpenalty=100
\ifydir \fi

However, we cannot use them under LuaTEX-ja. Instead of them, we have to write as the following.
\ltjsetparameter{kanjiskip=10pt} \dimen0=\ltjgetparameter{kanjiskip}
\ltjsetparameter{tbaselineshift=0.1\zw}
\dimen0=\ltjgetparameter{tbaselineshift}
\ltjsetparameter{prebreakpenalty={`,100}}
\ifnum\ltjgetparameter{direction}=4 \fi

Note that pTEX adds new two useful units, namely zw and zh. As shown above, they are changed to \zw and \zh respectively in LuaTEX-ja.\footnote{LuaTEX-ja 20200127.0 introduces \ltj@zw and \ltj@zh, which are copy of \zw and \zh.}

■ Linebreak after a Japanese character   In pTEX, a line break after Japanese character is ignored (and doesn’t yield a space), since line breaks (in source files) are permitted almost everywhere in Japanese texts. However, LuaTEX-ja doesn’t have this feature completely, because of a specification of LuaTeX. For the detail, see Section 14.
Spaces related to Japanese characters  The insertion process of glues/kerns between two Japanese characters and between a Japanese character and other characters (we refer glues/kerns of both kinds as JAglue) is rewritten from scratch.

- As LuaTeX’s internal ligature handling is node-based (e.g., ofice doesn’t prevent ligatures), the insertion process of JAglue is now node-based.
- Furthermore, nodes between two characters which have no effects in line break (e.g., \special node) and kerns from italic correction are ignored in the insertion process.
- Caution: due to above two points, many methods which did for the dividing the process of the insertion of JAglue in \TeX are not effective anymore. In concrete terms, the following two methods are not effective anymore:
  ちょっと ちょっと
  If you want to do so, please put an empty horizontal box (hbox) between it instead:
  ちょっと

- In the process, two Japanese fonts which only differ in their “real” fonts are identified.

Directions  From version 20150420.0, LuaTeX-ja supports vertical writing. We implement this feature by using callbacks of LuaTeX; so it must not be confused with Ω-style direction support of LuaTeX itself. Due to implementation, the dimension returned by \wd, \ht, or \dp depends on the content of the register only. This is major difference with \TeX.

\discretionary  Japanese characters in discretionary break (\discretionary) is not supported.

Greek and Cyrillic letters, and ISO 8859-1 symbols  By default, LuaTeX-ja uses Japanese fonts to typeset Greek and Cyrillic letters, To change this behavior, put \ltjsetparameter{jacharrange={-2,-3}} in the preamble. For the detailed description, see Subsection 4.1.

From version 20150906.0, characters which belongs both ISO 8859-1 and JIS X 0208, such as ¶ and §, are now typeset in alphabetic fonts.

1.3 Notations

In this document, the following terms and notations are used:

- Characters are classified into following two types. Note that the classification can be customized by a user (see Subsection 4.1).
  - JAchar: standing for characters which is used in Japanese typesetting, such as Hira-gana, Katakana, Kanji, and other Japanese punctuation marks.
  - ALchar: standing for all other characters like latin alphabets.

We say alphabetic fonts for fonts used in ALchar, and Japanese fonts for fonts used in JAchar.

- A word in a sans-serif font with underline (like prebreakpenalty) means an internal parameter for Japanese typesetting, and it is used as a key in \ltjsetparameter command.

- A word in a sans-serif font without underline (like fontspec) means a package or a class of \TeX.

- In this document, natural numbers start from zero. \omega denotes the set of all natural numbers which can be used in \TeX.
1.4 About the project

■ Project Wiki  Project Wiki is under construction.

- https://osdn.jp/projects/luatex-ja/wiki/FrontPage%28en%29 (English)

This project is hosted by OSDN.

■ Members

- Hironori KITAGAWA
- Yusuke KUROKI
- Tomoaki HONDA
- Kazuki MAEDA
- Noriyuki ABE
- Shuzaburo SAITO
- Takayuki YATO
- Munehiro YAMAMOTO
- MA Qiyuan
2 Getting Started

2.1 Installation

The following packages are needed for the Lua\TeX-ja package.

- Lua\TeX 1.10.0 (or later) (DVI output (outputmode=8 is not supported.)
- recent luatotfload (v3.1 or later recommended)
- adobemapping (Adobe cmap and pdfmapping files)
- E\TeX 2020-02-02 patch level 5 or later (if you want to use Lua\TeX-ja with E\TeX 2020-02-02)
- etoolbox (if you want to use Lua\TeX-ja with E\TeX 2020-02-02)
- everysec (only for E\TeX 2020-02-02 and 2020-10-01)
- filehook, atbegshi (only for E\TeX 2020-02-02)
- ltxcmds, pdftexcmds
- fontspec v2.7c (or later)
- Harano Aji fonts (https://github.com/trueroad/HaranoAjiFonts)
  More specifically, HaranoAjiMincho-Regular and HaranoAjiGothic-Medium.

Now Lua\TeX-ja is available from CTAN (in the macros/luatex/generic/luatexja directory), and the following distributions:

- \TeX Live (in texmf-dist/tex/luatex/luatexja)
- W32\TeX (in luatexja.tar.xz)
- Mi\TeX (in luatexja.tar.xz)

Harano Aji fonts are also available in these distributions (haranoaji in \TeX Live and Mi\TeX, and luatexja.tar.xz in W32\TeX).

HarfBuzz and Lua\TeX-ja

Using Lua\TeX-ja with LuaHB\TeX (Lua\TeX integrated with HarfBuzz) is not well tested. Maybe documents can typeset without an error, but with unwanted results (especially, vertical typesetting and \CID).

Especially, We don’t recommend defining a Japanese font with HarfBuzz, by specifying Renderer=Harfbuzz etc. (fontspec) or mode=harf (otherwise).

Manual installation

1. Download the source, by one of the following method. At the present, Lua\TeX-ja has no stable release.
   - Clone the Git repository by
     $ git clone git://git.osdn.jp/gitroot/luatex-ja/luatexja.git
   - Download the tar.gz archive of HEAD in the master branch from
     http://git.osdn.jp/view?p=luatex-ja/luatexja.git;a=snapshot;h=HEAD;sf=tgz.

Note that the master branch, and hence the archive in CTAN, are not updated frequently; the forefront of development is not the master branch.

2. Extract the archive. You will see src/ and several other sub-directories. But only the contents in src/ are needed to work Lua\TeX-ja.
3. If you downloaded this package from CTAN, you have to run following commands to generate classes:

```bash
$ cd src
$ lualatex ltjclasses.ins
$ lualatex ltjsclasses.ins
$ lualatex ltjjtxdoc.ins
```

4. Copy all the contents of `src/` into one of your `TEXMF` tree. `TEXMF/tex/luatex/luatexja/` is an example location. If you cloned entire Git repository, making a symbolic link of `src/` instead copying is also good.

5. If `mktexlar` is needed to update the file name database, make it so.

### 2.2 Cautions

For changes from p\TeX, see Subsection 1.2.

- The encoding of your source file must be UTF-8. Other encodings, such as EUC-JP or Shift-JIS, are not supported.

- Lua\TeX-ja is very slower than p\TeX, and uses a lot of memory.

- **(Outdated) note for MiKTeX users** Lua\TeX-ja requires that several CMap files\(^2\) must be found from Lua\TeX. Strictly speaking, those CMaps are needed only in the first run of Lua\TeX-ja after installing or updating. But it seems that MiKTeX does not satisfy this condition, so you will encounter an error like the following:

```
! LuaTeX error ...iles (x86)/MiKTeX 2.9/tex/luatex/luatexja/ltj-rmlgbm.lua
bad argument #1 to 'open' (string expected, got nil)
```

If so, please execute a batch file which is written on the Project Wiki (English). This batch file creates a temporary directory, copy CMaps in it, run a test file which loads Lua\TeX-ja in this directory, and finally delete the temporary directory.

- Note that when Lua\TeX-ja is loaded in plain Lua\TeX, we cannot use color specification on font loading, such as

```
\font\hoge=lmroman10-regular.otf:color=FF0000 % \font primitive
```

This is because codes for shifting baseline in math mode (Lua\TeX-ja) collide with and prevents loading codes for font color (luaotfload) in these environments. **We recommend to use \LaTeX 2020-02-02 (or later)**, since we can avoid this collision in there.

### 2.3 Using in plain \TeX

To use Lua\TeX-ja in plain \TeX, simply put the following at the beginning of the document:

```
\input luatexja.sty
```

This does minimal settings (like `ptex.tex`) for typesetting Japanese documents:

- The following 12 Japanese fonts are preloaded:

<table>
<thead>
<tr>
<th>direction</th>
<th>classification</th>
<th>font name</th>
<th>&quot;10 pt&quot;</th>
<th>&quot;7 pt&quot;</th>
<th>&quot;5 pt&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>yoko (horizontal)</td>
<td>mincho</td>
<td>HaranoAjiMincho-Regular</td>
<td>\tenmin</td>
<td>\sevenmin</td>
<td>\fivemin</td>
</tr>
<tr>
<td></td>
<td>gothic</td>
<td>HaranoAjiMincho-Medium</td>
<td>\tengt</td>
<td>\sevengt</td>
<td>\fivegt</td>
</tr>
<tr>
<td>tate (vertical)</td>
<td>mincho</td>
<td>HaranoAjiMincho-Regular</td>
<td>\tentmin</td>
<td>\seventmin</td>
<td>\fivetmin</td>
</tr>
<tr>
<td></td>
<td>gothic</td>
<td>HaranoAjiMincho-Medium</td>
<td>\tentgt</td>
<td>\seventgt</td>
<td>\fivegt</td>
</tr>
</tbody>
</table>

The "default" Japanese fonts (and JFMs for them) can be modified by defining \ltj@stdmcfont etc. before one inputs luatexja.sty (Subsection 7.3).

- A character in an alphabetic font is generally smaller than a Japanese font in the same size. So actual size specification of these Japanese fonts is in fact smaller than that of alphabetic fonts, namely scaled by 0.962216.

- The amount of glue that are inserted between a JAchar and an ALchar (the parameter xkanjiskip) is set to

\[
(0.25 \cdot 0.962216 \cdot 10)^{+1}_{-1} \text{pt} = 2.40554 \text{ pt}^{+1}_{-1} \text{pt}.
\]

2.4 Using in \LaTeX

Using in \LaTeX 2\kern-.1667em\raise.7ex\hbox{\scriptsize\bf e} is basically same. To set up the minimal environment for Japanese, you only have to load luatexja.sty:
\begin{verbatim}
\usepackage{luatexja}
\end{verbatim}

It also does minimal settings (counterparts in \TeX are plfonts.dtx and pldefs.ltx).

- Font encodings for Japanese fonts are JY3 (for horizontal direction) and JT3 (for vertical direction).

- Traditionally, Japanese documents use only two families: mincho (明朝体) and gothic (ゴシック体). 

\textit{mincho} is used in the main text, while \textit{gothic} is used in the headings or for emphasis.

Here \jttdefault specifies the Japanese font family in \verb or \verbatim environment, and its default value is \mcdefault (mincho family).\textsuperscript{3} \LaTeX-ja does not define commands to only switch current Japanese font family to \jttdefault.

- By default, the following fonts are used for these two families.

\begin{center}
\begin{tabular}{|l|c|c|c|}
\hline
\textbf{classification} & \textbf{family} & \textbf|\texttt\textmdseries| & \textbf|\texttt\textbfseries| & \textbf{scale} \\
\hline
\textit{mincho} (明朝体) & mc & HaranoAjiMincho-Regular & HaranoAjiGothic-Medium & 0.962216 \\
\textit{gothic} (ゴシック体) & gt & HaranoAjiGothic-Medium & HaranoAjiGothic-Medium & 0.962216 \\
\hline
\end{tabular}
\end{center}

- Note that the bold series (series \texttt bx or b) in both family are same as the medium series of gothic family. There is no italic nor slanted shape for these \texttt mc and \texttt gt.

- From version 20181102.0, one can specifies disablejfam option at loading \LaTeX-ja. This option prevents loading a patch for \LaTeX, which are needed to support Japanese characters in math mode.

Without disablejfam option, one can typeset Japanese characters in math mode as $\$\$ (see Page 11) as before. Japanese characters in math mode are typeset by the font family \texttt mc.

- If you use the beamer class with the default font theme (which uses sans serif fonts) and with \LaTeX-ja, you might want to change default Japanese fonts to the gothic family. The following line changes the default Japanese font family to gothic:
\begin{verbatim}
\renewcommand{\kanjifamilydefault}{\gtdefault}
\end{verbatim}

\textsuperscript{3} When \ltjsclasses classes are used, or luatexja-fontspec (or luatexja-preset) is loaded with match option, \jttfamily changes the current Japanese font family to \jttdefault. These classes and packages also redefine \jttdefault to \gtdefault (gothic family).
However, above settings are not sufficient for Japanese-based documents. To typeset Japanese-based documents, you are better to use class files other than article.cls, book.cls, and so on. At the present, Lua\TeX-ja has the counterparts of jclasses (standard classes in p\LaTeX) and jsclasses (classes by Haruhiko Okumura), namely, ltjclasses\footnote{ltjarticle.cls, ltjbook.cls, ltjreport.cls, ltjarticle.cls, ltjbook.cls, ltjreport.cls. The latter ltjt*.cls are for vertically written Japanese documents.} and ltjsclasses\footnote{ltjsarticle.cls, ltjsbook.cls, ltjsreport.cls, ltjsarticle.cls, ltjsbook.cls, ltjsreport.cls. The former jsclasses uses only \TeX code, but ltjsclasses uses Lua code.}.

Original jsclasses use \texttt{\mag} primitive to set the main document font size. However, Lua\TeX does not support \texttt{\mag} in PDF output, so ltjsclasses uses the \texttt{nomag*} option\footnote{Same effect as the BXjscls classes (by Takayuki Yato) and jsclasses. However, these classes uses only \TeX code, but ltjsclasses uses Lua code.} by default to set the main font size. If this causes some unexpected behavior, specify \texttt{nomag} option in \texttt{\documentclass}.

\section*{geometry package and classes for vertical writing}

It is well-known that the geometry package produces the following error, when classes for vertical writing is used:

\begin{verbatim}
! Incompatible direction list can't be boxed.
\begindvi \unvbox \begindvibox
  \global \let \begindvi \empty
\end{verbatim}

Now, Lua\TeX-ja automatically applies the patch \texttt{lltjp-geometry} to the geometry package, when the direction of the document is \texttt{tate} (vertical writing). This patch \texttt{lltjp-geometry} also can be used in p\LaTeX; for the detail, please refer \texttt{lltjp-geometry.pdf} (Japanese).

\section{Changing Fonts}

\subsection{plain \TeX and \LaTeXe}

\subsubsection{plain \TeX}
To change Japanese fonts in plain \TeX, you must use the command \texttt{\jfont} and \texttt{\tfont}. So please see Subsection 7.1.

\subsubsection{\LaTeXe (NFSS2)}
For \LaTeXe, Lua\TeX-ja adopted most of the font selection system of p\LaTeXe (in \texttt{plfonts.dtx}).

\begin{center}
\begin{tabular}{lllll}
\hline
 encoding & family & series & shape & selection \\
\hline
 Alphabetic fonts & \romanencoding & \romanfamily & \romanseries & \romanshape & \useroman \\
 Japanese fonts & \kanjicoding & \kanjifamily & \kanjiseries & \kanjishape & \usekanji \\
 both & \- & \- & \fontseries & \fontshape & \- \\
 auto select & \fontencoding & \fontfamily & \- & \- & \usefont \\
\hline
\end{tabular}
\end{center}

• \texttt{\fontfamily}, \texttt{\fontseries}, and \texttt{\fontshape} try to change attributes of Japanese fonts, as well as those of alphabetic fonts. Of course, \texttt{\selectfont} is needed to select current text fonts.

Note that \texttt{\fontshape} always changes current alphabetic font shape, but it does not change current Japanese font shape if the target shape is unavailable for current Japanese encoding/family/series. For the detail, see Subsection 10.2.

• \texttt{\fontencoding}\{\{encoding\}\} changes the encoding of alphabetic fonts or Japanese fonts depending on the argument. For example, \texttt{\fontencoding\{JY3\}} changes the encoding of Japanese fonts to JY3, and \texttt{\fontencoding\{T1\}} changes the encoding of alphabetic fonts to T1. \texttt{\fontfamily} also changes the current Japanese font family, the current alphabetic font family, or \texttt{both}. For the detail, see Subsection 10.2.

• For defining a Japanese font family, use \texttt{\DeclareKanjiFamily} instead of \texttt{\DeclareFontFamily}. (In previous version of Lua\TeX-ja, using \texttt{\DeclareFontFamily} didn’t cause any problem. But this no longer applies the current version.)

• Defining a Japanese font shape can be done by usual \texttt{\DeclareFontShape}:

\begin{verbatim}
\DeclareFontShape{JY3}{mc}{b}{n}{<-> s*HaranoAjiMincho-Bold:jfm=ujis;-kern}{% Harano Aji Mincho Bold
\end{verbatim}
Table 1. Commands of luatexja-fontspec

<table>
<thead>
<tr>
<th></th>
<th>\jfontspec</th>
<th>\setmainjfont</th>
<th>\setsansjfont</th>
<th>\setmonojfont</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese fonts</td>
<td>\newjfontfamily</td>
<td>\renewjfontfamily</td>
<td>\setjfontfamily</td>
<td></td>
</tr>
<tr>
<td>Alphabetic fonts</td>
<td>\newfamily</td>
<td>\renewfamily</td>
<td>\setfamily</td>
<td></td>
</tr>
<tr>
<td>Japanese fonts</td>
<td>\newjfontface</td>
<td>\defaultjfontfeatures</td>
<td>\addjfontfeatures</td>
<td></td>
</tr>
<tr>
<td>Alphabetic fonts</td>
<td>\newfontface</td>
<td>\defaultfontfeatures</td>
<td>\addfontfeatures</td>
<td></td>
</tr>
</tbody>
</table>

■Japanese characters in math mode Since \TeX supports Japanese characters in math mode, there are sources like the following:

1. $f_{\text{高温}}$~($f_{\text{high temperature}}$).
2. $y=(x-1)^2+2$ よって $y > 0$
3. $5 \in \text{素}: = \{ p \in \mathbb N : \text{$p$ is a prime} \}$.

We (the project members of Lua\TeX-ja) think that using Japanese characters in math mode are allowed if and only if these are used as identifiers. In this point of view,

- The lines 1 and 2 above are not correct, since “高温” in above is used as a textual label, and “いつも” is used as a conjunction.
- However, the line 3 is correct, since “素” is used as an identifier.

Hence, in our opinion, the above input should be corrected as:

1. $f_{\text{高温}}$.
2. $y = (x-1)^2 + 2$ よって $y > 0$
3. $5 \in \text{素}: = \{ p \in \mathbb N : p \text{ is a prime} \}$.

We also believe that using Japanese characters as identifiers is rare, hence we don’t describe how to change Japanese fonts in math mode in this chapter. For the method, please see Subsection 7.6.

When Lua\TeX-ja is loaded with disablejffam option, one cannot write Japanese characters in math mode as $\text{素}$. At that case, one have to use \textbf{mbox} (or \textit{text} in the amsmath package).

3.2 luatexja-fontspec package

To use the functionality of the fontspec package to Japanese fonts, it is needed to load the luatexja-fontspec package in the preamble, as follows:

\usepackage[(options)]{luatexja-fontspec}

This luatexja-fontspec package automatically loads luatexja and fontspec packages, if needed.

In the luatexja-fontspec package, several commands are defined as counterparts of original commands in the fontspec package (see Table 1):

The package option of luatexja-fontspec are the followings:

match

If this option is specified, usual family-changing commands such as \texttt{rmfamily}, \texttt{textrm}, \texttt{sffamily}, ... also change Japanese font family.

pass=(opts)

(Obsoleted) Specify options (opts) which will be passed to the fontspec package.
Override the ratio of the font size of Japanese fonts to that of alphabetic fonts. The default value is determined as follows:

- The value of `\Cjascale` is used, if this control sequence is already defined.
- It is calculated automatically from the current Japanese font at the loading of the package, if `\Cjascale` is not defined.

`\Cjascale` is defined in `ltjclasses` and `ltjsclasses`.

All other options listed above are simply passed to the `fontspec` package. This means that two lines below are equivalent, for example.

```
\usepackage[no-math]{fontspec}
\usepackage{luatexja-fontspec}
```

Note that kerning information in a font is not used (that is, kern feature is set off) by default in these seven (or eight) commands. This is because of the compatibility with previous versions of LuaTEX-ja (see 7.1).

Below is an example of `\jfontspec`.

```
\jfontspec[CJKShape=NLC]{HaranoAjiMincho-Regular}
```

3.3 Presets of Japanese fonts

With `luatexja-preset` package, one use one of "preset" to simplify Japanese font setting. For details of package options, and those of each presets, please see Subsection 12.6. The following presets are defined:

- haranoaji, hiragino-pro, hiragino-pron,ipa,ipa-hg,ipaex,ipaex-hg,kozuka-pró,kozuka-prón,kozuka-pro,moga-mobo,moga-mobo-ex,bizud,morisawa-prón,morisawa-pro,ms,ms-hg,noembed,noto-otc,noto-otf,sourcehan,sourcehan-jp,ume,yu-osx,yu-win,yu-win10

For example, this document loads `luatexja-preset` package by

```
\usepackage[haranoaji]{luatexja-preset}
```

which means that Harano Aji fonts will be used in this document.

3.4 `\CID`, `\UTF`, and macros in `japanese-otf` package

Under p\TeX, `japanese-otf` package (developed by Shuzaburo Saito) is used for typesetting characters which is in Adobe-Japan1-6 CID but not in JIS X 0208. Since this package is widely used, Lua\TeX-ja supports some of functions in the japanese-otf package, as an external package `luatexja-otf`.

4 Changing Internal Parameters

There are many internal parameters in Lua\TeX-ja. And due to the behavior of Lua\TeX, most of them are not stored as internal register of \TeX, but as an original storage system in Lua\TeX-ja. Hence, to assign or acquire those parameters, you have to use commands `\ltjsetparameter` and `\ltjgetparameter`. 
Table 2. Characters in predefined character range 8.

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>§ (U+00A7)</td>
<td>Section Sign</td>
</tr>
<tr>
<td>' (U+00B0)</td>
<td>Degree sign</td>
</tr>
<tr>
<td>' (U+00B4)</td>
<td>Spacing acute</td>
</tr>
<tr>
<td>× (U+00D7)</td>
<td>Multiplication sign</td>
</tr>
<tr>
<td>' (U+00B4)</td>
<td>Spacing acute</td>
</tr>
</tbody>
</table>

Table 3. Unicode blocks in predefined character range 1.

<table>
<thead>
<tr>
<th>Code</th>
<th>Block</th>
</tr>
</thead>
<tbody>
<tr>
<td>U+0080–U+00FF</td>
<td>Latin-1 Supplement</td>
</tr>
<tr>
<td>U+0100–U+017F</td>
<td>Latin Extended-A</td>
</tr>
<tr>
<td>U+0180–U+024F</td>
<td>Latin Extended-B</td>
</tr>
<tr>
<td>U+0250–U+02AF</td>
<td>IPA Extensions</td>
</tr>
<tr>
<td>U+02B0–U+02FF</td>
<td>Spacing Modifier Letters</td>
</tr>
<tr>
<td>U+1E00–U+1EFF</td>
<td>Latin Extended Additional</td>
</tr>
<tr>
<td>U+0300–U+036F</td>
<td>Combining Diacritical Marks</td>
</tr>
</tbody>
</table>

4.1 Range of JRefs

LuaTeX-ja divides the Unicode codespace U+0000–U+10FFFF into character ranges, numbered 1 to 217. The grouping can be (globally) customized by \texttt{\textbackslash \ltxdefcharrange}. The next line adds whole characters in Supplementary Ideographic Plane and the character "漢" to the character range 100.

\texttt{\textbackslash \ltxdefcharrange{100}{"20000-"2FFFF,`漢}}

A character can belong to only one character range. For example, whole SIP belong to the range 4 in the default setting of LuaTeX-ja, and if one executes the above line, then SIP will belong to the range 100 and be removed from the range 4.

The distinction between \texttt{ALchar} and \texttt{JAchar} is performed by character ranges. This can be edited by setting the jacharrange parameter. For example, the code below is just the default setting of LuaTeX-ja, and it sets

- a character which belongs character ranges 1, 4, 5, and 8 is \texttt{ALchar},
- a character which belongs character ranges 2, 3, 6, 7, and 9 is \texttt{JAchar}.

\texttt{\ltxsetparameter{jacharrange={-1, +2, +3, -4, -5, +6, +7, -8, +9}}}

The argument to jacharrange parameter is a list of non-zero integer. Negative integer \(n\) in the list means that "each character in the range \(n\) is an \texttt{ALchar}, and positive integer \(+n\) means that "... is a \texttt{JAchar}.

Note that characters U+0000–U+007F are always treated as an \texttt{ALchar} (this cannot be customized).

Default character ranges LuaTeX-ja predefines nine character ranges for convenience. They are determined from the following data:

- Blocks in Unicode 12.0.0.
- The Adobe-Japan1-UCS2 mapping between a CID Adobe-Japan1- and Unicode.
- The PXbase bundle for upTeX by Takayuki Yato.

Now we describe these nine ranges. The superscript "J" or "A" after the number shows whether each character in the range is treated as \texttt{JAchar} or not by default. These settings are similar to the \texttt{prefercjk} settings defined in PXbase bundle. Any characters equal to or above U+0080 which does not belong to these eight ranges belongs to the character range 217.

Range 8\(^A\) The intersection of the upper half of ISO 8859-1 (Latin-1 Supplement) and JIS X 0208 (a basic character set for Japanese). The character list is indicated in Table 2.

Range 1\(^A\) Latin characters that some of them are included in Adobe-Japan1-7. This range consists of the Unicode ranges indicated in Table 3, except characters in the range 8 above.

Range 2\(^J\) Greek and Cyrillic letters. JIS X 0208 (hence most of Japanese fonts) has some of these characters.
Table 4. Unicode blocks in predefined character range 3.

<table>
<thead>
<tr>
<th>Block</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>U+2070–U+209F</td>
<td>Superscripts and Subscripts</td>
</tr>
<tr>
<td>U+20A0–U+20CF</td>
<td>Currency Symbols</td>
</tr>
<tr>
<td>U+2100–U+214F</td>
<td>Letterlike Symbols</td>
</tr>
<tr>
<td>U+2150–U+218F</td>
<td>Number Forms</td>
</tr>
<tr>
<td>U+2190–U+21FF</td>
<td>Arrows</td>
</tr>
<tr>
<td>U+2200–U+22FF</td>
<td>Mathematical Operators</td>
</tr>
<tr>
<td>U+2210–U+221F</td>
<td>Comb. Diacritical Marks for Symbols</td>
</tr>
<tr>
<td>U+2300–U+23FF</td>
<td>Miscellaneous Technical</td>
</tr>
<tr>
<td>U+2500–U+257F</td>
<td>Box Drawing</td>
</tr>
<tr>
<td>U+2580–U+259F</td>
<td>Block Elements</td>
</tr>
<tr>
<td>U+25A0–U+25FF</td>
<td>Geometric Shapes</td>
</tr>
<tr>
<td>U+2600–U+26FF</td>
<td>Miscellaneous Symbols</td>
</tr>
<tr>
<td>U+2700–U+27BF</td>
<td>Dingbats</td>
</tr>
<tr>
<td>U+2900–U+297F</td>
<td>Supplemental Arrows-B</td>
</tr>
<tr>
<td>U+2980–U+29FF</td>
<td>Misc. Math Symbols-B</td>
</tr>
<tr>
<td>U+2B00–U+2BFF</td>
<td>Misc. Symbols and Arrows</td>
</tr>
</tbody>
</table>

Table 5. Characters in predefined character range 9.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>U+2002</td>
<td>En space</td>
</tr>
<tr>
<td>U+2010</td>
<td>Hyphen</td>
</tr>
<tr>
<td>U+2011</td>
<td>Non-breaking hyphen</td>
</tr>
<tr>
<td>U+2012</td>
<td>Em dash</td>
</tr>
<tr>
<td>U+2013</td>
<td>Horizontal bar</td>
</tr>
<tr>
<td>U+2014</td>
<td>Double vertical line</td>
</tr>
<tr>
<td>U+2015</td>
<td>Left single quotation mark</td>
</tr>
<tr>
<td>U+2016</td>
<td>Single low-9 quotation mark</td>
</tr>
<tr>
<td>U+2017</td>
<td>Right double quotation mark</td>
</tr>
<tr>
<td>U+2018</td>
<td>Right single quotation mark</td>
</tr>
<tr>
<td>U+2019</td>
<td>Double low-9 quotation mark</td>
</tr>
<tr>
<td>U+201A</td>
<td>Dagger</td>
</tr>
<tr>
<td>U+201B</td>
<td>Double dagger</td>
</tr>
<tr>
<td>U+201C</td>
<td>Bullet</td>
</tr>
<tr>
<td>U+201D</td>
<td>Em dash</td>
</tr>
<tr>
<td>U+201E</td>
<td>Horizontal ellipsis</td>
</tr>
<tr>
<td>U+201F</td>
<td>Single left-pointing angle quot.</td>
</tr>
<tr>
<td>U+2020</td>
<td>Reference mark</td>
</tr>
<tr>
<td>U+2021</td>
<td>Two dot leader</td>
</tr>
<tr>
<td>U+2022</td>
<td>Bullet</td>
</tr>
<tr>
<td>U+2023</td>
<td>Double prime</td>
</tr>
<tr>
<td>U+2024</td>
<td>Single right-pointing angle quot.</td>
</tr>
<tr>
<td>U+2025</td>
<td>Horizontal bar</td>
</tr>
<tr>
<td>U+2026</td>
<td>Single left-pointing angle quot.</td>
</tr>
<tr>
<td>U+2027</td>
<td>Asterism</td>
</tr>
<tr>
<td>U+2028</td>
<td>Two asterisks aligned vertically</td>
</tr>
<tr>
<td>U+2029</td>
<td>Overline</td>
</tr>
<tr>
<td>U+202A</td>
<td>Underline</td>
</tr>
<tr>
<td>U+202B</td>
<td>Double question mark</td>
</tr>
<tr>
<td>U+202C</td>
<td>Question exclamation mark</td>
</tr>
</tbody>
</table>

- **U+0370–U+03FF**: Greek and Coptic
- **U+0400–U+04FF**: Cyrillic
- **U+1F00–U+1FFF**: Greek Extended

**Range 3** Miscellaneous symbols. The block list is indicated in Table 4.

**Range 9** The intersection of the “General Punctuation” block (U+2000–U+206F) and Adobe-Japan1-7 character collection. This character range characters in Table 5.

**Range 4** Characters usually not in Japanese fonts. This range consists of almost all Unicode blocks which are not in other predefined ranges. Hence, instead of showing the block list, we put the definition of this range itself.

```latex
\texttt{\textbackslash jdefcharrange{4}{%}
"590-"10FF, "1200-"1DFF, "2440-"245F, "27C0-"28FF, "2A00-"2AFF,
"2B00-"2E7F, "4DC0-"4DFF, "A000-"A5F, "A980-"ABFF, "E000-"F8FF,
"F800-"FE0F, "FE20-"FE2F, "FE70-"FEFF, "10000-"1AFF, "1B170-"1F0FF,
"1F300-"1FFFF, ..., (and characters in U+2000–U+206F which are not in range 9)
}\} % non-Japanese
```

**Range 5** Surrogates and Supplementary Private Use Areas.

**Range 6** Characters used in Japanese. The block list is indicated in Table 6.

**Range 7** Characters used in CJK languages, but not included in Adobe-Japan1-7. The block list is indicated in Table 7.

**Notes on U+0080–U+00FF** You should treat characters in textttU+0080–U+00FF as \texttt{ALchar}, when you use traditional 8-bit fonts, such as the marvosym package.
4.2 **kanjiskip** and **xkanjiskip**

**JAglue** is divided into the following three categories:

- Glues/kerns specified in JFM. If \texttt{\protect\textbackslash inhibit\textbackslash glue} is issued around a **JChar**, this glue will not be inserted at the place.
- The default glue which inserted between two **JChars** (**kanjiskip**).
- The default glue which inserted between a **JChar** and an **AlChar** (**xkanjiskip**).

The value (a skip) of **kanjiskip** or **xkanjiskip** can be changed as the following. Note that only their values at the end of a paragraph or a hbox are adopted in the whole paragraph or the whole hbox.

\texttt{\protect\textbackslash ltjsetparameter\{kanjiskip\={0pt plus 0.4pt minus 0.4pt}, xkanjiskip\={0.25}\zw plus 1pt minus 1pt\}}

Here \texttt{\zw} is a internal dimension which stores fullwidth of the current Japanese font. This \texttt{\zw} can be used as the unit \texttt{zw} in \texttt{\Pi X}.

The value of these parameter can be get by \texttt{\protect\textbackslash ltjgetparameter}. Note that the result by \texttt{\protect\textbackslash ltjgetparameter} is not the internal quantities, but a string (hence \texttt{\the} cannot be prefixed).

\texttt{kanjiskip: \protect\textbackslash ltjgetparameter\{kanjiskip\}, xkanjiskip: \protect\textbackslash ltjgetparameter\{xkanjiskip\}}

It may occur that JFM contains the data of “ideal width of **kanjiskip**” and/or “ideal width of **xkanjiskip**”. To use these data from JFM, set the value of **kanjiskip** or **xkanjiskip** to \texttt{\maxdimen} (these “ideal width” cannot be retrived by \texttt{\protect\textbackslash ltjgetparameter}).

15
4.3 Insertion setting of \texttt{xkanjiskip}

It is not desirable that \texttt{xkanjiskip} is inserted into every boundary between \texttt{JAchars} and \texttt{ALchars}. For example, \texttt{xkanjiskip} should not be inserted after opening parenthesis (e.g., compare “(あ)” and “(あ)”). Lua\TeX{}-ja can control whether \texttt{xkanjiskip} can be inserted before/after a character, by changing \texttt{jaxspmode} for \texttt{JAchars} and \texttt{alxspmode} parameters \texttt{ALchars} respectively.

\begin{verbatim}
\ltjsetparameter{jaxspmode={``あ',preonly},
alxspmode={``!,postonly}}
pあqいう
\end{verbatim}

The second argument \texttt{preonly} means that the insertion of \texttt{xkanjiskip} is allowed before this character, but not after. the other possible values are \texttt{postonly}, \texttt{allow}, and \texttt{inhibit}.

\texttt{jaxspmode} and \texttt{alxspmode} use a same table to store the parameters on the current version. Therefore, line 1 in the code above can be rewritten as follows:

\begin{verbatim}
\ltjsetparameter{alxspmode={``あ',preonly}, jaxspmode={'\!',postonly}}
\end{verbatim}

One can use also numbers to specify these two parameters (see Subsection 8.1).

If you want to enable/disable all insertions of \texttt{kanjiskip} and \texttt{xkanjiskip}, set \texttt{autospacing} and \texttt{autoxspacing} parameters to \texttt{true}/\texttt{false}, respectively.

4.4 Shifting the baseline

To make a match between a Japanese font and an alphabetic font, sometimes shifting of the baseline of one of the pair is needed. In p\TeX{}, this is achieved by setting \texttt{\baselineshift} (or \texttt{\baselineshift}) to a non-zero length (the baseline of \texttt{ALchar} is shifted below). However, for documents whose main language is not Japanese, it is good to shift the baseline of Japanese fonts, but not that of alphabetic fonts. Because of this, Lua\TeX{}-ja can independently set the shifting amount of the baseline of alphabetic fonts and that of Japanese fonts.

\begin{center}
\begin{tabular}{|l|l|}
\hline
Alphabetic fonts & yalbaselineshift parameter & talbaselineshift parameter \\
Japanese fonts & yjabaselineshift parameter & tjabaselineshift parameter \\
\hline
\end{tabular}
\end{center}

Here the horizontal line in the below example is the baseline of a line.

\begin{verbatim}
\vrule width 150pt height 0.2pt depth 0.2pt \
hskip-120pt
\ltjsetparameter{yjabaselineshift=0pt, yalbaselineshift=0pt}abc
\ltjsetparameter{yjabaselineshift=5pt, yalbaselineshift=2pt}abc
\end{verbatim}

There is an interesting side-effect: characters in different size can be vertically aligned center in a line, by setting two parameters appropriately. The following is an example (beware the value is not well tuned):

\begin{verbatim}
\vrule width 150pt height4.417pt depth-4.217pt \\
kern-150pt
\large \texttt{xyz漢字}
\{\scriptsize
\ltjsetparameter{yjabaselineshift=-1.757pt, yalbaselineshift=-1.757pt}漢字xyzあいう
\}
\end{verbatim}

Note that setting positive \texttt{yalbaselineshift} or \texttt{talbaselineshift} parameters does not increase the depth of one-letter syllable \texttt{p} of \texttt{Alchar}, if its left-protrusion (\texttt{\lpcode}) and right-protrusion (\texttt{\rpcode}) are both non-zero. This is because

- These two parameters are implemented by setting \texttt{yoffset} field of a glyph node, and this does not increase the depth of the glyph.
• To cope with the above situation, LuaTeX-ja automatically supplies a rule in every syllable.

• However, we cannot use this “supplying a rule” method if a syllable comprises just one letter whose \lpcode and \rpcode are both non-zero.

This problem does not apply for \vjabaselineshift nor \tjabaselineshift, because a \JAchar is encapsulated by a horizontal box if needed.

4.5 kinsoku parameters and OpenType features

Among parameters which related to Japanese word-wrapping process (kinsoku shori),

\jaxspmode, \alxspmode, \prebreakpenalty, \postbreakpenalty and \kcatcode

are stored by each character codes.

OpenType font features are ignored in these parameters. For example, a fullwidth katakana “ア” on line 10 in the below input is replaced to its halfwidth variant “ア”, by \hwid feature. However, the penalty inserted after it is 10 which is the \postbreakpenalty of “ア”, not 20.

\begin{verbatim}
\jsetparameter{postbreakpenalty={ア, 10}}
\jsetparameter{postbreakpenalty={ア, 20}}
\newcommand\showpostpena[1]{%
  \leavevmode\setbox0=\hbox{#1\hbox{}}%}
\showpostpena{ア}, \showpostpena{ア}, \showpostpena{ア}
\addjfontfeatures{CharacterWidth=Half}\showpostpena{ア}
\end{verbatim}
Part II

Reference

5 \texttt{\textbackslash catcode in Lua\TeX\-ja}

5.1 Preliminaries: \texttt{\textbackslash kcatcode in \texttt{p\TeX} and \texttt{up\TeX}}

In \texttt{p\TeX} and \texttt{up\TeX}, the value of \texttt{\textbackslash kcatcode} determines whether a Japanese character can be used in a control word. For the detail, see Table 8.

\texttt{\textbackslash kcatcode} can be set by a row of JIS X 0208 in \texttt{p\TeX}, and generally by a Unicode block\footnote{\texttt{up\TeX} divides U+FF00–U+FDEF (Halfwidth and Fullwidth Forms) into three subblocks, and \texttt{\textbackslash kcatcode} can be set by a subblock.} in \texttt{up\TeX}. So characters which can be used in a control word slightly differ between \texttt{p\TeX} and \texttt{up\TeX}.

5.2 Case of Lua\TeX\-ja

The role of \texttt{\textbackslash kcatcode} in \texttt{p\TeX} and \texttt{up\TeX} can be divided into the following four kinds, and Lua\TeX\-ja can control these four kinds separately:

\begin{itemize}
  \item Distinction between \texttt{JAchar} or \texttt{ALchar} is controlled by the character range, see Subsection 4.1.
  \item Whether the character can be used in a control word is controlled by setting \texttt{\textbackslash kcatcode} to 11 (enabled) or 12 (disabled), as usual.
  \item Whether \texttt{icharwidowpenalty} can be inserted before the character is controlled by the lowermost bit of the \texttt{\textbackslash kcatcode} parameter.
  \item Linebreak after a \texttt{JAchar} does not produce a space.
\end{itemize}

Default setting of \texttt{\textbackslash kcatcode} of Unicode characters are located in

\texttt{plain Lua\TeX} \texttt{luatex-unicode-letters.tex}, which is based on \texttt{unicode-letters.tex} (for \TeX).

\texttt{Lua\TeX} now included in \texttt{\texttt{p\TeX} kernel as unicode-letters.def}.

However, the default setting of \texttt{\textbackslash kcatcode} differs between \texttt{Xe\TeX} and Lua\TeX, by the following reasons:

\begin{itemize}
  \item (plain format) \texttt{luatex-unicode-letters.tex} is based on old \texttt{unicode-letters.tex}.
  \item The latter half of \texttt{unicode-letters.tex} and \texttt{unicode-letters.def} sets\texttt{\textbackslash kcatcode} of several characters to 11, via setting \texttt{\textbackslash XeTeXCharClass}. However, this latter half does not exist (plain case), or not executed (\texttt{\texttt{p\TeX} case) in Lua\TeX}.
\end{itemize}

In other words,

\texttt{plain Lua\TeX} Kanji nor kana characters cannot be used in a control word, in the default setting of plain Lua\TeX.

\texttt{Lua\TeX} In recent (2015-10-01 or later) Lua\TeX, Kanji and kana characters in a control word is supported (these catcode are 11), but not fullwidth alphanumerics and several other characters.

This would be inconvenient for \texttt{p\TeX} users to shifting to Lua\TeX\-ja, since several control words containing Kanji or other fullwidth characters, such as \texttt{\textbackslash x年月日} or \texttt{1年目西暦} are used in \texttt{p\TeX}. Hence, Lua\TeX\-ja have a counterpart of \texttt{unicode-letters.tex} for Lua\TeX, to match the \texttt{\textbackslash kcatcode setting} with that of \texttt{Xe\TeX}.

5.3 Non-kanji characters in a control word

Because the engine differ, so non-kanji JIS X 0208 characters which can be used in a control word differ in \texttt{p\TeX}, in \texttt{up\TeX}, and in Lua\TeX\-ja. Table 9 shows the difference. Except for four characters “・” “・” “・” “・”, Lua\TeX\-ja admits more characters in a control word than \texttt{up\TeX}.

Difference becomes larger, if we consider non-kanji JIS X 0213 characters. For the detail, see \url{https://github.com/h-kitagawa/kct}. 

\begin{table}
\end{table}
6 Directions

LuaTEX supports four Ω-style directions: TLT, TRT, RTT and LTL. However, neither directions are not well-suited for typesetting Japanese vertically, hence we implemented vertical writing by rotating TLT-box by 90 degrees.

LuaTEX-ja supports four directions, as shown in Table 10. The second column (yoko direction) is just horizontal writing, and the third column (tate direction) is vertical writing. The fourth column (dtou direction) is actually a hidden feature of pTEX. We implemented this for debugging purpose. The fifth column (utod direction) corresponds the “tate (math) direction” of pTEX.

Directions can be changed by `\yoko`, `\tate`, `\dtou`, `\utod`, only when the current list is null. These commands cannot be executed in unrestricted horizontal modes, nor math modes. The direction of a math formula is changed to `\utod`, when the direction outside the math formula is `\tate` (vertical writing).

6.1 Boxes in different direction

As in pTEX, one can use boxes of different direction in one document. The below is an example.

```
\begin{verbatim}
\text{ここは横組\% yoko}
\hbox{\tate\% tate}
\hbox{縦組\% tate}

\hbox{の中に\% tate}
\hbox{横組の内容\% yoko}
\hbox{を挿入する}\}
\text{また横組に戻る\% yoko}
\end{verbatim}
```

Table 8. \texttt{\kcatcode} in upTEX

<table>
<thead>
<tr>
<th>\kcatcode</th>
<th>meaning</th>
<th>control word</th>
<th>widow penalty</th>
<th>linebreak</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>non-cjk</td>
<td>(treated as usual \TeX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>kanji</td>
<td>Y</td>
<td>Y</td>
<td>ignored</td>
</tr>
<tr>
<td>17</td>
<td>kana</td>
<td>Y</td>
<td>Y</td>
<td>ignored</td>
</tr>
<tr>
<td>18</td>
<td>other</td>
<td>N</td>
<td>N</td>
<td>ignored</td>
</tr>
<tr>
<td>19</td>
<td>hangul</td>
<td>Y</td>
<td>Y</td>
<td>space</td>
</tr>
</tbody>
</table>

Table 9. Difference of the set of non-kanji JIS X 0208 characters which can be used in a control word

<table>
<thead>
<tr>
<th>row col.</th>
<th>p\TeX</th>
<th>up\TeX</th>
<th>LuaTEX-ja</th>
</tr>
</thead>
<tbody>
<tr>
<td>(U+30FB)</td>
<td>1</td>
<td>6</td>
<td>N</td>
</tr>
<tr>
<td>(U+309B)</td>
<td>1</td>
<td>11</td>
<td>N</td>
</tr>
<tr>
<td>(U+309C)</td>
<td>1</td>
<td>12</td>
<td>Y</td>
</tr>
<tr>
<td>(U+FF48)</td>
<td>1</td>
<td>14</td>
<td>N</td>
</tr>
<tr>
<td>(U+FF3E)</td>
<td>1</td>
<td>16</td>
<td>N</td>
</tr>
<tr>
<td>(U+FF3F)</td>
<td>1</td>
<td>18</td>
<td>N</td>
</tr>
<tr>
<td>(U+3008)</td>
<td>1</td>
<td>23</td>
<td>N</td>
</tr>
<tr>
<td>(U+4EDB)</td>
<td>1</td>
<td>24</td>
<td>N</td>
</tr>
<tr>
<td>(U+3005)</td>
<td>1</td>
<td>25</td>
<td>N</td>
</tr>
<tr>
<td>(U+3006)</td>
<td>1</td>
<td>26</td>
<td>N</td>
</tr>
<tr>
<td>(U+3007)</td>
<td>1</td>
<td>27</td>
<td>N</td>
</tr>
<tr>
<td>(U+30FC)</td>
<td>1</td>
<td>28</td>
<td>N</td>
</tr>
<tr>
<td>(U+FF8F)</td>
<td>1</td>
<td>31</td>
<td>N</td>
</tr>
<tr>
<td>(U+FF3C)</td>
<td>1</td>
<td>32</td>
<td>N</td>
</tr>
</tbody>
</table>

\text{Greek letters (row 6)} \ Y \ N \ Y
\text{Cyrillic letters (row 7)} \ N \ N \ Y
Table 10. Directions supported by Lua\TeX-ja

<table>
<thead>
<tr>
<th>Commands</th>
<th>horizontal (\yoko direction)</th>
<th>vertical (\tate direction)</th>
<th>dtou direction</th>
<th>utod direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning of the page</td>
<td>\yoko</td>
<td>\tate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beginning of the line</td>
<td>\tate</td>
<td>\left</td>
<td>\right</td>
<td></td>
</tr>
<tr>
<td>Used Japanese font</td>
<td>\yoko \underbar</td>
<td>vertical (\tate \underbar)</td>
<td>\yoko \underbar \right \underbar</td>
<td>\right \underbar</td>
</tr>
<tr>
<td>Example</td>
<td>\textcolor{red}{\textbullet}</td>
<td>\textcolor{red}{\textbullet}</td>
<td>\textcolor{red}{\textbullet}</td>
<td>\textcolor{red}{\textbullet}</td>
</tr>
</tbody>
</table>

(Notation used in $\Omega$)

TLT RTR, RTT LBL RTR

Table 11 shows how a box is arranged when the direction inside the box and that outside the box differ.

\textbf{\texttt{\textbackslash w}d and \texttt{\textbackslash d}irection} In p\TeX, \texttt{\textbackslash w}, \texttt{\textbackslash d}, \texttt{\textbackslash p} means the dimensions of a box register \textit{with respect to the current direction}. This means that the value of \texttt{\textbackslash w}\texttt{\textbackslash d} etc. might differ when the current direction is different, even if \texttt{\textbackslash box} stores the same box. However, this no longer applies in Lua\TeX-ja.

1 \setbox0=\hbox to 20pt{foo} 
2 \the\wd0, \hbox{\tate\vrule\the\wd0} 
3 \wd0=100pt 
4 \the\wd0, \hbox{\tate \the\wd0} 

To access box dimensions \textit{with respect to current direction}, one has to use the following commands instead of \texttt{\textbackslash w}d wtc.

\begin{verbatim}
\newcommand{\ltjgetwd}{\texttt{\textbackslash ltjgetwd}(\texttt{\textbackslash num})} 
\newcommand{\ltjgetht}{\texttt{\textbackslash ltjgetht}(\texttt{\textbackslash num})} 
\newcommand{\ltjgetdp}{\texttt{\textbackslash ltjgetdp}(\texttt{\textbackslash num})}
\end{verbatim}

These commands return an \textit{internal dimension} of \texttt{\textbackslash box}(\texttt{\textbackslash num}) with respect to the current direction. One can use these in \texttt{\textbackslash dimexpr} primitive, as the followings.

\begin{verbatim}
\dimexpr 2\ltjgetwd42-3pt\relax, \the\ltjgetwd1701
\end{verbatim}

The following is an example.

\begin{verbatim}
\setparindent0pt 
\setbox32767=\hbox{\yoko \kyokum} 
\vbox{\hsize=20mm \yoko YOKO \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.} 
\vbox{\hsize=20mm \tate TATE \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.} 
\vbox{\hsize=20mm \dtou DTOU \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.} 
\vbox{\hsize=20mm \utod UTOD \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.} 
\end{verbatim}

\begin{verbatim}
\ltjsetwd(\texttt{\textbackslash num})=\langle\texttt{dimen}\rangle, \ltjsetht(\texttt{\textbackslash num})=\langle\texttt{dimen}\rangle, \ltjsetdp(\texttt{\textbackslash num})=\langle\texttt{dimen}\rangle 
\end{verbatim}

These commands set the dimension of \texttt{\textbackslash box}(\texttt{\textbackslash num}). One does not need to group the argument \texttt{\textbackslash num}; four calls of \texttt{\ltjsetwd} below have the same meaning.

\begin{verbatim}
\ltjsetwd42 20pt, \ltjsetwd42=20pt, \ltjsetwd42=20pt, \ltjsetwd42=20pt
\end{verbatim}

6.2 Getting current direction

The \texttt{\textbackslash direction} parameter returns the current direction, and the \texttt{\textbackslash boxdir} parameter (with the argument \texttt{\textbackslash num}) returns the direction of a box register \texttt{\textbackslash box}(\texttt{\textbackslash num}). The returned value of these parameters are a \texttt{string}:
Table 11. Boxes in different direction

<table>
<thead>
<tr>
<th>Direction</th>
<th>yoko</th>
<th>tate</th>
<th>dtou</th>
<th>utod</th>
<th>(empty)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returned value</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

### 6.3 Overridden box primitives

To cope with multiple directions, the following primitives are overridden by LuaTeX-ja, using \protect{def}.

\leavevmode\def{DIR}{\ltjgetparameter{direction}}
\unhbox{\yoko{DIR}}, \unbox{\tate{DIR}},
\unhbox{\dtou{DIR}}, \unbox{\utod{DIR}},
\unvbox{\tate{DIR}}, \unvbox{\utod{DIR}},
\unvbox{\tate math: \DIR}$\|$ \unvbox{\utod math: \DIR}$\$
\setbox2=\hbox{\tate{DIR}}\ltjgetparameter{boxdir}{2}

\vadjust{\langle material \rangle}
\insert{\langle number \rangle}{\langle material \rangle}
\lastbox
\raise{\langle dimen \rangle}{\langle box \rangle} etc., \vcenter
\vcenter
7 Font Metric and Japanese Font

7.1 \jfont

To load a font as a Japanese font (for horizontal direction), you must use the \jfont instead of \font, while \jfont admits the same syntax used in \font. LuaTeX-ja automatically loads luaotfload package, so TrueType/OpenType fonts with features can be used for Japanese fonts:

\[\text{\jfont}\langle\text{jfont cs}\rangle=\ldots; jfm=\langle\text{JFM name}\rangle; \ldots; jfmvar=\langle\text{identifier}\rangle; \ldots\]

\langle JFM name \rangle The name of a (horizontal) JFM. LuaTeX-ja searches and loads jfm-⟨JFM name⟩.lua.

\langle JFM features \rangle An optional comma-separated list of JFM options. Enclosing braces ({} are optional, but this does not escape any characters. The contents of this list can be accessed by a table luatexja.jfont.jfm feature from a JFM, at its loading. See Figure 2 for an example.

Note that any JFM files which is shipped with LuaTeX-ja does not use this feature.

\langle identifier \rangle An optional string.

LuaTeX-ja “does not distinguish” two Japanese fonts which uses same JFM and are the same size. Here “uses same JFM” means that all of ⟨JFM name⟩, ⟨JFM features⟩ and ⟨identifier⟩ of two fonts agree.

For example, The first “】” and “】 in Figure 1 are typeset in different real fonts. However, because they use the same JFMs and their size are same, LuaTeX-ja inserts penalties, glues and kerns as if these two character are typeset in a same font. Namely, the glue between these characters is halfwidth, as in “】】”. However, this does not applies with \F and \H in Figure 1, because their ⟨identifier⟩ are different.

■ Horizontal JFMs The following horizontal JFMs are shipped with LuaTeX-ja.

jfm-ujis.lua A standard horizontal JFM of LuaTeX-ja. This file is based on upm1minr-h.tfm, a metric for UTF/OTF package that is used in upTeX. When you are going to use the luatexja-otf package, you should use this JFM.
\A: (nil) \B: [kana] = true, [kern] = "0.5", [ps] = false, \C: [down] = "0.2", [kern] = "0.5", \D: [down] = "0.2", [kern] = "0.5",

\(1\) \small{\ltjsetparameter{differentjfm=both}\tabcolsep=.5\zw\)
\(2\)% \printjfmfeat is defined in the source of this document
\(3\)\jfont\A=HaranoAjiMincho-Regular:jfm=testf at 9pt \printjfmfeat\A
\(4\)\jfont\B=HaranoAjiMincho-Bold:jfm=testf/kern=0.5,-ps,+kana at 9pt \printjfmfeat\B
\(5\)\jfont\C=HaranoAjiGothic-Regular:jfm=testf/kern=0.5,down=0.2 at 9pt \printjfmfeat\C
\(6\)\jfont\D=HaranoAjiGothic-Bold:jfm=testf/down=0.2,kern=0.5 at 9pt \printjfmfeat\D
\(7\)\def\TEST#1{\string#1&{#1あ漢イ字}{\Aイ字}{\Bイ字}{\Cイ字}{\Dイ字}}
\(8\)\vspace{-4\baselineskip}\hfill\ttfamily
\(9\)\begin{tabular}{lllll}
\(10\)&\string\A&\string\B&\string\C&\string\D\
\TEST\A\
\TEST\B\
\TEST\C\
\TEST\D\
\end{tabular}
\(11\)% No space between "漢" and "イ" iff two Japanese fonts use same JFM
\(12\)\ltjsetparameter{differentjfm=paverage}

Figure 2. Example of JFM features

Table 12. Differences between horizontal JFMs shipped with \LaTeX-ja

<table>
<thead>
<tr>
<th>jfm-jis.lua</th>
<th>jfm-ujis.lua</th>
<th>jfm-min.lua</th>
</tr>
</thead>
<tbody>
<tr>
<td>A counterpart for jis.tfm, &quot;JIS font metric&quot; which is widely used in \TeX. A major difference between jfm-ujis.lua and this jfm-jis.lua is that most characters under jfm-ujis.lua are square-shaped, while that under jfm-jis.lua are horizontal rectangles.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A counterpart for min10.tfm, which is one of the default Japanese font metric shipped with \TeX.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A JFM for proportional typesetting. This JFM doesn’t have any information of character dimension (width, height, depth), nor glues/kerns information.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Another JFM for proportional typesetting. In contrast to jfm-prop.lua, this JFM has informations of character height and depth.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

See Table 12 for the difference among jfm-ujis.lua, jfm-jis.lua, jfm-min.lua.

■Using kerning information in a font Some fonts have information for inter-glyph spacing. \LaTeX-ja 20140324.0 or later treats kerning spaces like an italic correction; any glue and/or kern from the JFM and a kerning space can coexist. See Figure 3 for detail.

Note that in \setmainfont etc. which are provided by luatexja-fontspec package, kerning option is set off (Kerning=0ff) by default, because of the compatibility with previous versions of \LaTeX-ja.

23
\begin{multicols}{2}
\begin{verbatim}
\newcommand\test{\vrule ダ イ ナ ミ ッ ク ダ イ ク マ \vrule}
\end{verbatim}
\end{multicols}

Figure 3. Kerning information and \texttt{kanjiskip}

\begin{verbatim}
\leavevmode\ltjsetparameter{kanjiskip=0pt plus 3\zw}
\vrule\hbox to 15\zw{あ「い」う,えお}\vrule
\jfont\E=HaranoAjiMincho-Regular:extend=1.5;jfm=ujis;-kern
\vrule\hbox to 15\zw{あ「い」う,えお}\vrule
\end{verbatim}

Figure 4. \texttt{ltjksp} \textit{feature}

\textbf{extend and slant} The following setting can be specified as OpenType font features:

\begin{verbatim}
extend=⟨extend⟩ expand the font horizontally by ⟨extend⟩.
slant=⟨slant⟩ slant the font.
\end{verbatim}

Note that Lua\TeX-ja doesn’t adjust JFMs by these \texttt{extend} and \texttt{slant} settings; one have to write new JFMs on purpose. For example, the following example uses the standard JFM \texttt{jfm-ujis.lua}, hence the letterspacing and the width of italic corrections are not correct:

\begin{verbatim}
\jfont\E=HaranoAjiMincho-Regular:extend=1.5;jfm=ujis;-kern
\jfont\S=HaranoAjiMincho-Regular:slant=1;jfm=ujis;-kern
\end{verbatim}

\textbf{ltjksp \textit{feature}} \texttt{kanjiskip\_natural}, \texttt{kanjiskip\_stretch}, \texttt{kanjiskip\_shrink} keys (Page ???) makes the Lua\TeX-ja inserts not only a glue which is specified by a JFM, and also the natural width/stretch part/shrink part of \texttt{kanjiskip}. This functionality is disabled by \texttt{-ltjksp} specification, as shown in Figure 4.

7.2 \texttt{tfont}

\texttt{tfont} loads a font as a Japanese font for vertical direction. This command admits the same syntax as in \texttt{font} and \texttt{jfont}. A font defined by \texttt{tfont} differs the following points from that by \texttt{jfont}:
Figure 5. \texttt{jpotf} "feature"

- OpenType Feature \texttt{vrt2} is automatically activated, unless \texttt{vert} and/or \texttt{vrt2} features are explicitly activated or deactivated (as the second line in the example below).

- Sometimes \texttt{vert} and/or \texttt{vrt2} are not activated while one specified activation of these feature. This is because the font does not define these features in current combination of script tag and language system identifier.

- Furthermore, a glyph is automatically rotated 90 degrees, if it is not replaced by \texttt{vert} feature for any script for any language, and if it is marked as ‘r’ or “Tr” in UAX #50.

- \texttt{jfm} uses a vertical JFM instead of a horizontal JFM. \LaTeX-ja ships following vertical JFMs:
  \begin{itemize}
  \item \texttt{jfm-ujisv.lua} A standard vertical JFM in \LaTeX-ja. This JFM is based on \texttt{upnmlminr-v.tfm}, a metric for UTF/OTF package that is used in up\TeX.
  \item \texttt{jfm-tmin.lua} A counterpart for \texttt{tmin10.tfm}, which is one of the default Japanese font metric shipped with p\TeX.
  \end{itemize}

- If \texttt{vert} and/or \texttt{vrt2} features are activated, one can specify \texttt{jpotf} to additional substitutions. By default, it substitutes ideographic comma/period for fullwidth comma/period, and double prime quotation marks for double quotation marks (See Figure 5). One can customize substitutions by lua function \texttt{luatexja.jfont.register_vert_replace} (see Japanese version of this manual).

\section{Default Japanese fonts and JFMs}

If following commands are defined at loading \LaTeX-ja package, these change default Japanese fonts and JFMs for them:

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{\ltj@stdmcfont}</td>
<td>The default Japanese font for the mincho family.</td>
</tr>
<tr>
<td>\texttt{\ltj@stdgtfont}</td>
<td>The default Japanese font for the gothic family.</td>
</tr>
<tr>
<td>\texttt{\ltj@stdyokojfm}</td>
<td>The default JFM for horizontal direction.</td>
</tr>
</tbody>
</table>

\footnote{If the font does not define the \texttt{vrt2} feature, \texttt{vert} is used instead.}
The default JFM for vertical direction.

For example,
\def\ltj@stdmcfont{IPAMincho}
\def\ltj@stdgtfont{IPA Gothic}

makes that IPA Mincho and IPA Gothic will be used as default Japanese fonts, instead of Harano Aji fonts.

This feature is intended for classes which use special JFMs\(^9\). It is recommended to use \texttt{\luatexja-preset} or \texttt{\luatexja-fontspec} package to select standard fonts in ordinary \LaTeX sources.

For compatibility with earlier versions, Lua\TeX-ja reads \texttt{luatexja.cfg} automatically if it is found by Lua\TeX. One should not overuse this \texttt{luatexja.cfg}; it will overwrite the definition of \texttt{\ltj@stdmcfont} and others.

### 7.4 Prefix \texttt{psft}

Besides “\texttt{file}” and “\texttt{name}” prefixes which are introduced in the \texttt{luaotfload} package, Lua\TeX-ja adds “\texttt{psft}” prefix in \texttt{\jfont} (and \texttt{\font}), to specify a “name-only” Japanese font which will not be embedded to PDF. Note that these non-embedded fonts under current Lua\TeX has Identity-H encoding, and this violates the standard ISO32000-1:2008 ([10]).

*OpenType font features, such as “+jp90”, have no meaning in name-only fonts using “\texttt{psft}” prefix, because we can’t expect what fonts are actually used by the PDF reader.* Note that extend and slant settings (see above) are supported with \texttt{psft} prefix, because they are only simple linear transformations.

#### \texttt{cid key}

The default font defined by using \texttt{psft} prefix is for Japanese typesetting; it is Adobe-Japan1-7 CID-keyed font. One can specify \texttt{cid} key to use other CID-keyed non-embedded fonts for Chinese or Korean typesetting.

\begin{verbatim}
\jfont\testJ={psft:Ryumin-Light:cid=Adobe-Japan1-7;jfm=jis}% Japanese
\jfont\testD={psft:Ryumin-Light:jfm=jis}% default: Adobe-Japan1-7
\jfont\testC={psft:AdobeMingStd-Light:cid=Adobe-CNS1-7;jfm=jis}% Traditional Chinese
\jfont\testG={psft:SimSun:cid=Adobe-GB1-5;jfm=jis}% Simplified Chinese
\jfont\testK={psft:Batang:cid=Adobe-Korea1-2;jfm=jis}% Korean
\jfont\testKR={psft:SourceHanSerifAKR9:cid=Adobe-KR-9;jfm=jis}% Korean
\end{verbatim}

Note that the code above specifies \texttt{jfm-jis.lua}, which is for Japanese fonts, as JFM for Chinese and Korean fonts.

At present, Lua\TeX-ja supports only 5 values written in the sample code above. Specifying other values, e.g.,

\begin{verbatim}
\jfont\testJ={psft:Ryumin-Light:cid=Adobe-Japan2;jfm=jis}
\end{verbatim}

produces the following error:

\begin{verbatim}
! Package luatexja Error: bad cid key 'Adobe-Japan2'.
See the luatexja package documentation for explanation.
<Type H <return> for immediate help.
<to be read again>
\par
1.78
? h
I couldn't find any non-embedded font information for the CID 'Adobe-Japan2'. For now, I'll use 'Adobe-Japan1-6'. Please contact the LuaTeX-ja project team.
\end{verbatim}

\(^9\)This is because commands has @ in their names.
### 7.5 Structure of a JFM file

A JFM file is a Lua script which has only one function call:

```lua
luatexja.jfont.define_jfm { ... }
```

Real data are stored in the table which indicated above by `{ ... }`. So, the rest of this subsection are devoted to describe the structure of this table. Note that all lengths in a JFM file are floating-point numbers in design-size unit.

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>version</code></td>
<td>⟨version⟩ (optional, default value is 1) The version JFM. Currently 1, 2, and, 3 are supported</td>
</tr>
<tr>
<td><code>dir</code></td>
<td>⟨direction⟩ (required) The direction of JFM. ‘yoko’ (horizontal) or ‘tate’ (vertical) are supported.</td>
</tr>
<tr>
<td><code>zw</code></td>
<td>⟨length⟩ (required) The amount of the length of the “full-width”.</td>
</tr>
<tr>
<td><code>zh</code></td>
<td>⟨length⟩ (required) The amount of the “full-height” (height + depth).</td>
</tr>
<tr>
<td><code>kanjiskip</code></td>
<td>⟨natural⟩, ⟨stretch⟩, ⟨shrink⟩} (optional) This field specifies the “ideal” amount of <code>kanjiskip</code>. As noted in Subsection 4.2, if the parameter <code>kanjiskip</code> is <code>\maxdimen</code>, the value specified in this field is actually used (if this field is not specified in JFM, it is regarded as 0 pt). Note that ⟨stretch⟩ and ⟨shrink⟩ fields are in design-size unit too.</td>
</tr>
<tr>
<td><code>xkanjiskip</code></td>
<td>⟨natural⟩, ⟨stretch⟩, ⟨shrink⟩} (optional) Like the <code>kanjiskip</code> field, this field specifies the “ideal” amount of <code>xkanjiskip</code>.</td>
</tr>
<tr>
<td><code>chars</code></td>
<td>⟨character⟩, ...} (required except character class 0) This field is a list of characters which are in this character type $i$. This field is optional if $i = 0$, since all JAchar which do not belong any character classes other than 0 are in the character class 0 (hence, the character class 0 contains most of JAchars). In the list, character(s) can be specified in the following form:</td>
</tr>
<tr>
<td></td>
<td>- a Unicode code point</td>
</tr>
<tr>
<td></td>
<td>- the character itself (as a Lua string, like ‘あ’)</td>
</tr>
<tr>
<td></td>
<td>- a string like ‘あ*’ (the character followed by an asterisk)</td>
</tr>
<tr>
<td></td>
<td>- several “imaginary” characters (We will describe these later.)</td>
</tr>
<tr>
<td><code>width</code></td>
<td>⟨length⟩, <code>height=⟨length⟩</code>, <code>depth=⟨length⟩</code>, <code>italic=⟨length⟩</code>} (required) Specify the width of characters in character class $i$, the height, the depth and the amount of italic correction. All characters in character class $i$ are regarded that its width, height, and depth are as values of these fields. The default values are shown in Table 13.</td>
</tr>
</tbody>
</table>

![Table 13. Default values of width field and other fields](image-url)
Consider a Japanese character node which belongs to a character class whose the align field is 'middle'.

- The black rectangle is the imaginary body of the node. Its width, height, and depth are specified by JFM.
- Since the align field is 'middle', the “real” glyph is centered horizontally (the green rectangle) first.
- Furthermore, the glyph is shifted according to values of fields left and down. The ultimate position of the real glyph is indicated by the red rectangle.

![Figure 6. The position of the real glyph (horizontal Japanese fonts)](image)

(***)

![Figure 7. The position of the real glyph (vertical Japanese fonts)](image)

left={length}, down={length}, align={align}

These fields are for adjusting the position of the “real” glyph. Legal values of align field are 'left', 'middle', and 'right'. If one of these 3 fields are omitted, left and down are treated as 0, and align field is treated as 'left'. The effects of these 3 fields are indicated in Figure 6 and Figure 7.

In most cases, left and down fields are 0, while it is not uncommon that the align field is 'middle' or 'right'. For example, setting the align field to 'right' is practically needed when the current character class is the class for opening delimiters'.

kern={\[j\]=\{kern\}, \[j’\]=\{kern\}, \[ratio=\{ratio\}]}, ...}
glue={\[i\]=\{(width), (stretch), (shrink), [ratio=\{ratio\}, ...]}}, ...}

Specifies the amount of kern or glue which will be inserted between characters in character class i and those in character class j.

⟨ratio⟩ specifies how much the glue is originated in the “right” character. It is a real number between 0 and 1, and treated as 0.5 if omitted. For example, The width of a glue between an ideographic full stop “。” and a fullwidth middle dot “・” is three-fourth of fullwidth, namely halfwidth from the ideographic full stop, and quarter-width from the fullwidth middle dot. In this case, we specify ⟨ratio⟩ to 0.25/(0.5 + 0.25) = 1/3.

In case of glue, one can specify following additional keys in each \[j] subtable:

priority=⟨priority⟩ An integer in [−4, 3] (treated as 0 if omitted), or a pair of these integers {⟨stretch⟩, ⟨shrink⟩} (version 2 or later). This is used only in line adjustment with priority by luatexja-adjust (see Subsection 12.3). Higher value means the glue is easy to stretch, and is also easy to shrink.

kanjiskip_natural=(num), kanjiskip_stretch=(num), kanjiskip_shrink=(num)

These keys specifies the amount of the natural width of kanjiskip (the stretch/shrink part, respectively) which will be inserted in addition to the original JFM glue. Default values of them are all 0.

As an example, in jfm-ujis.lua, the standard JFM in horizontal writing, we have
• Between an ordinal letter “あ” and an ideographic opening bracket, we have a glue whose natural part and shrink part are both half-width, while its stretch part is zero. However, this glue also can be stretched as much as the stretch part of \texttt{kanjiskip} times the value of \texttt{kanjiskip\_stretch\_key} (1 in this case).

• Between an ideographic closing brackets (the ideographic comma “,” is included) and an ordinal letter, we have the same glue. Again, this glue also can be stretched as much as the stretch part of \texttt{kanjiskip} times the value of \texttt{kanjiskip\_stretch\_key} (1 in this case).

• Between an ideographic opening bracket and an ordinal letter and between an ordinal letter and an ideographic closing bracket, we have a glue whose natural part and stretch part are both zero, while its shrink part as much as the shrink part of \texttt{kanjiskip}.

Hence we have the following result:

\begin{verbatim}
\leavevmode
\setparameter{kanjiskip=0pt plus 3zw}
\vrule\hbox to 15zw{あ「い」う，えお}\vrule
\vrule\hbox{あ「い」う，えお}\vrule\par
\vrule\hbox to 6.5zw{あ「い」う，えお}\vrule
\end{verbatim}

end\_stretch=⟨kern⟩, end\_shrink=⟨kern⟩ (optional, version 1 only)
end\_adjust=⟨⟨kern⟩, ⟨kern⟩, ⟨kern⟩⟩ (optional, version 2 or later)
round\_threshold=⟨float⟩ (optional, version 3 or later, only available in character class 0)

\section*{Character to character classes}

We explain how the character class of a character is determined, using \texttt{jfm\_test.lua} which contains the following:

\begin{verbatim}
[0] = { chars = { '漢' }, align = 'left', left = 0.0, down = 0.0, width = 1.0, height = 0.88, depth = 0.12, italic=0.0, },
[2000] = { chars = { '、', '‘' }, align = 'left', left = 0.0, down = 0.0, width = 0.5, height = 0.88, depth = 0.12, italic=0.0, },
\end{verbatim}

Now consider the following input/output:

\begin{verbatim}
jfont\a=IPAexMincho:jfm=test;+hwid
\setbox0\hbox{\a ヒ\漢}\the\wd0
\end{verbatim}

Now we look why the above source outputs 15 pt.

1. The character “ヒ” is converted to its half width form “ヒ” by \texttt{hwid} feature.
2. According to the JFM, the character class of “ヒ” is 2000, hence its width is halfwidth.
3. The character class of “漢” is zero, hence its width is fullwidth.
4. Hence the width of \texttt{\hbox} equals to 15 pt.

This example shows that the character class of a character is generally determined after applying font features by luaotfload.

However, if the class determined by the glyph after application of features is zero, Lua\TeX-ja adopts the class determined by the glyph before application of features. The following input is an example.

\begin{verbatim}
jfont\a=HaranoAjiMincho-Regular:jfm=test;+vert
\a 漢。\inhibitglue 漢
\end{verbatim}
Table 14. Commands for Japanese math fonts

<table>
<thead>
<tr>
<th>Japanese fonts</th>
<th>alphabetic fonts</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\jfam \in [0, 256)$</td>
<td>$\fam$</td>
</tr>
<tr>
<td>jatextfont{${\jfam}, {\font_cs}$}</td>
<td>$\textfont(\fam)=$(${\font_cs}$)</td>
</tr>
<tr>
<td>jascriptfont{${\jfam}, {\font_cs}$}</td>
<td>$\scriptfont(\fam)=$(${\font_cs}$)</td>
</tr>
<tr>
<td>jaiscriptscriptfont{${\jfam}, {\font_cs}$}</td>
<td>$\scriptscriptfont(\fam)=$(${\font_cs}$)</td>
</tr>
</tbody>
</table>

Here, the character class of the ideographic full stop “。” (U+3002) is determined as follows:

1. As the case of “ヒ”, the ideographic full stop “。” is converted to its vertical form “圆满 (U+FE12) by vert feature.

2. The character class of “圆满”, according to the JFM is zero.

3. However, LuaTeX-ja remembers that this “圆满” is obtained from “。” by font features. The character class of “。” is non-zero value, namely, 2000.

4. Hence the ideographic full stop “。” in above belongs the character class 2000.

**Imaginary characters** As described before, one can specify several “imaginary characters” in chars field. The most of these characters are regarded as the characters of class 0 in \( \text{TeX} \). As a result, LuaTeX-ja can control typesetting finer than \( \text{TeX} \). The following is the list of imaginary characters:

- ‘boxbdd’
  The beginning/ending of a hbox, and the beginning of a noindented (i.e., began by \texttt{\noindent}) paragraph. If a hbox \( b \) begins (resp. ends) a glue or kern between this “charater” and a JChar, JAgue won’t be inserted before(resp. after) the hbox \( b \). \texttt{kanjiskip} and \texttt{xkanjiskip} around a hbox.

- ‘parbdd’
  The beginning of an (indented) paragraph.

- ‘jcharbdd’
  A boundary between JChar and anything else.

- ‘alchar’, ‘nox_alchar’
  (version 3 or later) A boundary between JChar and ALchar.

- ‘glue’
  (version 3 or later) A boundary between JChar, and, a glue or kern.

- \(-1\)
  The left/right boundary of an inline math formula.

**Porting JFM from \( \text{TeX} \)** See Japanese version of this manual.

### 7.6 Math font family

\( \text{TeX} \) handles fonts in math formulas by 16 font families\(^{10}\), and each family has three fonts: $\textfont$, $\scriptfont$ and $\scriptscriptfont$.

LuaTeX-ja’s handling of Japanese fonts in math formulas is similar; Table 14 shows counterparts to \( \text{TeX} \)’s primitives for math font families. There is no relation between the value of $\fam$ and that of $\jfam$; with appropriate settings, one can set both $\fam$ and $\jfam$ to the same value. Here $\{\font_cs\}$ in the argument of \texttt{jatextfont} etc. is a control sequence which is defined by $\jfam$, i.e., a horizontal Japanese font.

\(^{10}\)Omega, Aleph, Lua\( \text{TeX} \) and \( \varepsilon \)-up \( \text{TeX} \) can handle 256 families, but an external package is needed to support this in plain \( \text{TeX} \) and \( \text{LaTeX} \).
7.7 Callbacks

LuaTeX-ja also has several callbacks. These callbacks can be accessed via luatexbase.add_to_callback function and so on, as other callbacks.

luatexja.load_jfm callback
With this callback, one can overwrite JFMs. This callback is called when a new JFM is loaded.

```lua
function (<table> jfm_info, <string> jfm_name)
    return <table> new_jfm_info
end
```

The argument jfm_info contains a table similar to the table in a JFM file, except this argument has chars field which contains character codes whose character class is not 0.

An example of this callback is the 1tjarticle class, with forcefully assigning character class 0 to 'parbdd' in the JFM jfm-min.lua.

luatexja.define_jfont callback
This callback and the next callback form a pair, and you can assign characters which do not have fixed code points in Unicode to non-zero character classes. This luatexja.define_font callback is called just when new Japanese font is loaded.

```lua
function (<table> jfont_info, <number> font_number)
    return <table> new_jfont_info
end
```

jfont_info has the following fields, which may not overwritten by a user:

- **size** The font size specified at \texttt{\jfont} in scaled points (\(1\text{ sp} = 2^{-16}\text{ pt}\)).
- **zw, zh, kanjiskip, xkanjiskip** These are scaled value of those specified by the JFM, by the font size.
- **jfm** The internal number of the JFM.
- **var** The value of jfmvar key, which is specified at \texttt{\jfont}. The default value is the empty string.
- **chars** The mapping table from character codes to its character classes.
  
  The specification \texttt{[i].chars={⟨character⟩, ...}} in the JFM will be stored in this field as \texttt{chars={⟨⟨character⟩⟩= i, ...}.

- **char_type** For \(i \in \omega\), char_type[\(i\)] is information of characters whose class is \(i\), and has the following fields:
  
  - width, height, depth, italic, down, left are just scaled value of those specified by the JFM, by the font size.
  - align is a number which is determined from align field in the JFM:

  \[
  \begin{align*}
  1 & \text{ ('right' in JFM),} \\
  0.5 & \text{ ('middle' in JFM),} \\
  0 & \text{(otherwise).}
  \end{align*}
  \]

  For \(i, j \in \omega\), char_type[\(i\)][\(j\)] stores a kern or a glue which will be inserted between character class \(i\) and class \(j\).

  The returned table new_jfont_info also should include these fields, but you are free to add more fields (to use them in the luatexja.find_char_class callback). The font_number is a font number.

  A good example of this and the next callbacks is the luatexja-otf package, supporting "AJ1-xxx" form for Adobe-Japan1 CID characters in a JFM. This callback doesn’t replace any code of LuaTeX-ja.

luatexja.find_char_class callback
This callback is called just when LuaTeX-ja is trying to determine which character class a character chr_code belongs. A function used in this callback should be in the following form:
function (>number< char_class, >table< jfont_info, >number< chr_code)  
if char_class~=0 then return char_class  
else  

....  
return (>number< new_char_class or 0)  
end  
end

The argument char_class is the result of LuaTeX-ja’s default routine or previous function calls in this callback, hence this argument may not be 0. Moreover, the returned new_char_class should be as same as char_class when char_class is not 0, otherwise you will overwrite the LuaTeX-ja’s default routine.

\textbf{luatexja.set_width callback}

This callback is called when LuaTeX-ja is trying to encapsulate a \texttt{JAchar glyph_node}, to adjust its dimension and position.

\begin{verbatim}
function (>table< shift_info, >table< jfont_info, >table< char_type)  
return >table< new_shift_info  
end
\end{verbatim}

The argument shift_info and the returned new_shift_info have down and left fields, which are the amount of shifting down/left the character in a scaled point.

A good example is test/valign.lua. After loading this file, the vertical position of glyphs is automatically adjusted; the ratio (height : depth) of glyphs is adjusted to be that of letters in the character class 0. For example, suppose that

- The setting of the JFM: (height) = 88x, (depth) = 12x (the standard values of Japanese OpenType fonts);
- The value of the real font: (height) = 28y, (depth) = 5y (the standard values of Japanese TrueType fonts).

Then, the position of glyphs is shifted up by

\[
\frac{88x}{88x + 12x} (28y + 5y) - 28y = \frac{26}{25} y = 1.04y.
\]

8 Parameters

8.1 \texttt{\textbackslash ltjsetparameter}

As described before, \texttt{\textbackslash ltjsetparameter} and \texttt{\textbackslash ltjglobalsetparameter} are commands for accessing most parameters of LuaTeX-ja. One of the main reason that LuaTeX-ja didn’t adopted the syntax similar to that of \texttt{\textbackslash presetpenalty} (e.g., \texttt{\textbackslash prebreakpenalty} =10000) is the position of hpack_filter callback in the source of LuaTeX, see Section 13.

\texttt{\textbackslash ltjsetparameter} and \texttt{\textbackslash ltjglobalsetparameter} are commands for assigning parameters. These take one argument which is a key-value list. The difference between these two commands is the scope of assignment; \texttt{\textbackslash ltjsetparameter} does a local assignment and \texttt{\textbackslash ltjglobalsetparameter} does a global one by default. They also obey the value of \texttt{\textbackslash globaldefs}, like other assignments.

The following is the list of parameters which can be specified by the \texttt{\textbackslash ltjsetparameter} command. [\texttt{\textbackslash ces}] indicates the counterpart in \texttt{\textbackslash presetpenalty}, and symbols beside each parameter has the following meaning:

- “∗” : values at the end of a paragraph or a hbox are adopted in the whole paragraph or the whole hbox.
- “†” : assignments are always global.

\texttt{\textbackslash jcharwidowpenalty}(penalty)∗ [\texttt{\textbackslash jcharwidowpenalty}]

Penalty value for suppressing orphans. This penalty is inserted just after the last \texttt{JAchar} which is not regarded as a (Japanese) punctuation mark.
\texttt{kcattcode} = \{\langle \text{chr\_code} \rangle, \langle \text{natural number} \rangle \}\}^* \\
An additional attributes which each character whose character code is \langle \text{chr\_code} \rangle has. At version 20120506.0 or later, the lowermost bit of \langle \text{natural number} \rangle indicates whether the character is considered as a punctuation mark (see the description of \texttt{jcharwidowpenalty} above).

\texttt{prebreakpenalty} = \{\langle \text{chr\_code} \rangle, \langle \text{penalty} \rangle \}\}^* [\texttt{prebreakpenalty}]

Set a penalty which is inserted automatically before the character \langle \text{chr\_code} \rangle, to prevent a line starts from this character. For example, a line cannot started with one of closing brackets “)” , so Lua\TeX-ja sets \\
\texttt{\textbackslash ltxsetparameter\{prebreakpenalty=\{``,10000\}\}} by default.

\texttt{p\TeX} has following restrictions on \texttt{\textbackslash prebreakpenalty} and \texttt{\textbackslash postbreakpenalty}, but they don't exist in Lua\TeX-ja:

\begin{itemize}
  \item Both \texttt{\textbackslash prebreakpenalty} and \texttt{\textbackslash postbreakpenalty} cannot be set for the same character.
  \item We can set \texttt{\textbackslash prebreakpenalty} and \texttt{\textbackslash postbreakpenalty} up to 256 characters.
\end{itemize}

\texttt{\textbackslash postbreakpenalty} = \{\langle \text{chr\_code} \rangle, \langle \text{penalty} \rangle \}\}^* [\texttt{\textbackslash postbreakpenalty}]

Set a penalty which is inserted automatically after the character \langle \text{chr\_code} \rangle, to prevent a line ends with this character.

\texttt{\textbackslash jatextfont} = \{\langle \text{jfam} \rangle, \langle \text{jfont\_cs} \rangle \}\}^* [\texttt{\textbackslash textfont in \TeX}]

\texttt{\textbackslash javascriptfont} = \{\langle \text{jfam} \rangle, \langle \text{jfont\_cs} \rangle \}\}^* [\texttt{\textbackslash scriptfont in \TeX}]

\texttt{\textbackslash javascriptscriptfont} = \{\langle \text{jfam} \rangle, \langle \text{jfont\_cs} \rangle \}\}^* [\texttt{\textbackslash scriptscriptfont in \TeX}]

\texttt{yjabaselineshift} = \langle \text{dimen} \rangle

\texttt{yalbaselineshift} = \langle \text{dimen} \rangle [\texttt{\ybaselineshift}]

\texttt{tjabaselineshift} = \langle \text{dimen} \rangle

\texttt{talbaselineshift} = \langle \text{dimen} \rangle [\texttt{\tbaselineshift}]

\texttt{\textbackslash jaxspmode} = \{\langle \text{chr\_code} \rangle, \langle \text{mode} \rangle \}\}^* [\texttt{\textbackslash xspmode}]

Set whether inserting \texttt{xkanjiskip} is allowed before/after a \texttt{J\textbackslash A\textbackslash char} whose character code is \langle \text{chr\_code} \rangle. The followings are allowed for \langle \text{mode} \rangle:

\begin{itemize}
  \item \texttt{0, inhibit} Insertion of \texttt{xkanjiskip} is inhibited before the character, nor after the character.
  \item \texttt{1, preonly} Insertion of \texttt{xkanjiskip} is allowed before the character, but not after.
  \item \texttt{2, postonly} Insertion of \texttt{xkanjiskip} is allowed after the character, but not before.
  \item \texttt{3, allow} Insertion of \texttt{xkanjiskip} is allowed both before the character and after the character. This is the default value.
\end{itemize}

This parameter is similar to the \texttt{\inhibitxspcode} primitive of \texttt{p\TeX}, but not compatible with \texttt{\inhibitxspcode}.

\texttt{\textbackslash alxspmode} = \{\langle \text{chr\_code} \rangle, \langle \text{mode} \rangle \}\}^* [\texttt{\textbackslash xspmode}]

Set whether inserting \texttt{xkanjiskip} is allowed before/after a \texttt{A\textbackslash L\textbackslash char} whose character code is \langle \text{chr\_code} \rangle. The followings are allowed for \langle \text{mode} \rangle:

\begin{itemize}
  \item \texttt{0, inhibit} Insertion of \texttt{xkanjiskip} is inhibited before the character, nor after the character.
  \item \texttt{1, preonly} Insertion of \texttt{xkanjiskip} is allowed before the character, but not after.
  \item \texttt{2, postonly} Insertion of \texttt{xkanjiskip} is allowed after the character, but not before.
  \item \texttt{3, allow} Insertion of \texttt{xkanjiskip} is allowed before the character and after the character. This is the default value.
\end{itemize}

33
Note that parameters \texttt{jaxspmode} and \texttt{alxspmode} share a common table, hence these two parameters are synonyms of each other.

\texttt{autospacing}=(\texttt{bool}) \texttt{[\texttt{autospacing}]}

\texttt{autoxspacing}=(\texttt{bool}) \texttt{[\texttt{autoxspacing}]}

\texttt{kanjiskip}=(\texttt{skip})* \texttt{[\texttt{kanjiskip}]}

The default glue which inserted between two \texttt{JAchar}s. Changing current Japanese font does not alter this parameter, as \texttt{pTeX}.

If the natural width of this parameter is \texttt{maxdimen}, \texttt{LuaTeX-ja} uses the value which is specified in the JFM for current Japanese font (See Subsection 7.5).

\texttt{xkanjiskip}=(\texttt{skip})* \texttt{[\texttt{xkanjiskip}]}

The default glue which inserted between a \texttt{JAchar} and an \texttt{ALchar}. Changing current font does not alter this parameter, as \texttt{pTeX}.

As \texttt{kanjiskip}, if the natural width of this parameter is \texttt{maxdimen}, \texttt{LuaTeX-ja} uses the value which is specified in the JFM for current Japanese font (See Subsection 7.5).

\texttt{differentjfm}=(\texttt{mode})\dagger

Specify how glues/kerns between two \texttt{JAchar}s whose JFM (or size) are different. The allowed arguments are the followings:

\begin{itemize}
  \item average
  \item both
  \item large
  \item small
  \item pleft
  \item pright
  \item paverage
\end{itemize}

The default value is \texttt{paverage}.

\texttt{jacharrange}=(\texttt{ranges})

\texttt{kansujichar}={\{\texttt{digit}}, \langle \texttt{chr\_code}\rangle\}*[\texttt{kansujichar}]

\texttt{direction}=(\texttt{dir})\texttt{(always local)}

Assigning to this parameter has the same effect as \texttt{\yoko} (if \langle \texttt{dir}\rangle = 4), \texttt{\tate} (if \langle \texttt{dir}\rangle = 3), \texttt{\dtou} (if \langle \texttt{dir}\rangle = 1) or \texttt{\utod} (if \langle \texttt{dir}\rangle = 11). If the argument \langle \texttt{dir}\rangle is not one of 4, 3, 1 nor 11, the behavior of this assignment is undefined.

8.2 \texttt{\stringltjgetparameter}

\texttt{\stringltjgetparameter} is a control sequence for acquiring parameters. It always takes a parameter name as first argument.

\begin{itemize}
  \item \texttt{\stringltjgetparameter{differentjfm}},
  \item \texttt{\stringltjgetparameter{autospacing}},
  \item \texttt{\stringltjgetparameter{kanjiskip}},
  \item \texttt{\stringltjgetparameter{prebreakpenalty}}{}``
\end{itemize}

\texttt{The return value of \texttt{\stringltjgetparameter} is always a string, which is outputted by \texttt{tex.write()}. Hence any character other than space “ “ (U+0020) has the category code 12 (other), while the space has 10 (space).}

\begin{itemize}
  \item If first argument is one of the following, no additional argument is needed.
    \begin{itemize}
      \item jcharwidowpenalty
      \item yjabaselineshift
      \item yalbaselineshift
      \item autospacing
      \item autoxspacing
      \item kanjiskip
      \item xkanjiskip
      \item differentjfm
      \item direction
    \end{itemize}
    Note that \texttt{\stringltjgetparameter{autospacing}} and \texttt{\stringltjgetparameter{autoxspacing}} returns 1 or 0, not true nor false.
    \item If first argument is one of the following, an additional argument—a character code, for example—is needed.
    \begin{itemize}
      \item kcatcode
      \item prebreakpenalty
      \item postbreakpenalty
      \item jaxspmode
      \item alxspmode
    \end{itemize}
\end{itemize}

34
\texttt{\textbackslash \lt jgetparameter\{jasxpmode\}\{\ldots\}} and \texttt{\textbackslash \lt jgetparameter\{alxpmode\}\{\ldots\}} returns 0, 1, 2, or 3, instead of \texttt{\preonly} etc.

- \texttt{\lt jgetparameter\{jacharrange\}\{\texttt{range}\}} returns 0, 1, 2, or 3, instead of \texttt{\preonly} etc.

- For an integer \texttt{(digit)} between 0 and 9, \texttt{\lt jgetparameter\{kansujichar\}\{(digit)\}} returns the character code of the result of \texttt{\kansuji\{digit\}}.

- \texttt{\lt jgetparameter\{adjustdir\}} returns a integer which represents the direction of the surrounding vertical list. As \texttt{direction}, the return value 1 means \texttt{dou} direction, 3 means \texttt{tate} direction (vertical typesetting), and 4 means \texttt{yoko} direction (horizontal typesetting).

- For an integer \texttt{(reg\_num)} between 0 and 65535, \texttt{\lt jgetparameter\{boxdir\}\{(reg\_num)\}} returns the direction of \texttt{\box\{reg\_num\}}. If this box register is void, the returned value is zero.

- The following parameter names \textbf{cannot be specified} in \texttt{\lt jgetparameter}.

  \begin{itemize}
  \item \texttt{jatextfont, jascriptfont, jascriptscriptfont, jacharrange}
  \end{itemize}

  \texttt{\lt jgetparameter\{chartorange\}\{\texttt{chr\_code}\}} returns the range number which \texttt{\{chr\_code\}} belongs to (although there is no parameter named \texttt{chartorange}).

  If \texttt{(chr\_code)} is between 0 and 127, this \texttt{(chr\_code)} does not belong to any character range. In this case, \texttt{\lt jgetparameter\{chartorange\}\{(chr\_code)\}} returns \texttt{-1}.

  Hence, one can know whether \texttt{(chr\_code)} is \texttt{J\textbf{A}\textbf{c}har} or not by the following:

  \begin{verbatim}
  \texttt{\lt jgetparameter\{jacharrange\}\{\lt jgetparameter\{chartorange\}\{\texttt{chr\_code}\}\}}
  \end{verbatim}

  \begin{equation}
  \% 0 \text{ if } \text{JAchar}, 1 \text{ if } \text{ALchar}
  \end{equation}

- Because the returned value is string, the following conditionals do not work if \texttt{kanjiskip} (or \texttt{xkanjiskip}) has the stretch part or the shrink part.

  \begin{verbatim}
  \newif\if\tempskipa
  \ifdim\lt jgetparameter\{kanjiskip\}\v@ ... \fi
  \ifdim\lt jgetparameter\{xkanjiskip\}\v@ ... \fi
  \end{verbatim}

  The correct way is using a temporary register.

  \begin{verbatim}
  \tempskipa=\lt jgetparameter\{kanjiskip\} \ifdim\tempskipa\v@ ... \fi
  \tempskipa=\lt jgetparameter\{xkanjiskip\} \ifdim\tempskipa\v@ ... \fi
  \end{verbatim}

### 8.3 Alternative Commands to \texttt{\lt jsetparameter}

The basic method to set parameters of \LaTeX\-ja is to use \texttt{\lt jsetparameter} or \texttt{\lt jglobalsetparameter}. However, these commands are slow, because they parse a key-value list, so several alternative commands are used in \LaTeX\-ja. \textit{This subsection is not for general \LaTeX\-ja users.}

#### Setting \texttt{kanjiskip} or \texttt{xkanjiskip}

In \texttt{ltjclasses}, every size-changing command such as \texttt{\Large} changes \texttt{kanjiskip} and \texttt{xkanjiskip}. But a simple implementation, as the code below, is slow since two key-value lists are parsed by \texttt{\lt jsetparameter}:

\begin{verbatim}
\texttt{\lt jsetparameter\{kanjiskip\}0\zw plus .1\zw minus .01\zw}
\texttt{\tempskipa=\lt jgetparameter\{xkanjiskip\}}
\ifdim\tempskipa\v@ \else\fi
\texttt{\lt jsetparameter\{xkanjiskip\}0.1em}
\else\fi
\end{verbatim}

Hence, \LaTeX\-ja defines more primitive commands, namely \texttt{\lt j@setpar@global}, \texttt{\lt jsetkanjiskip}, and \texttt{\lt jsetxkanjiskip}. Here
and \ltjsetkanjiskip{10pt} has the same effect. The actual code of \ltjsclasses is shown below:

\begin{verbatim}
\ltj@setpar@global
\ltjsetkanjiskip{\z@ plus .1\zw minus .01\zw}
@tempskipa=\ltjgetparameter{xkanjiskip}
\ifdim@tempskipa>\z@
  \if@slide
    \ltjsetxkanjiskip.1em
  \else
    \ltjsetxkanjiskip.25em plus .15em minus .06em
  \fi
\fi
\end{verbatim}

Note that using \ltjsetkanjiskip or \ltjsetxkanjiskip alone, that is, without executing \ltj@setpar@global in advance, is not supported.

9 Other Commands for plain \TeX{} and \LaTeX{} 2ε

9.1 Commands for compatibility with \pif\TeX{}

The following commands are implemented for compatibility with \pif\TeX{}. Note that the former five commands don't support JIS X 0213, but only JIS X 0208. The last \kansuji converts an integer into its Chinese numerals.

\begin{verbatim}
\kuten, \jis, \euc, \sjis, \ucs, \kansuji
\end{verbatim}

These six commands take an internal integer, and returns a string.

\begin{verbatim}
\newcount\hoge
\hoge=2423\% \kansuji\hoge
\jis\hoge, \char\jis\hoge\kansuji1701
\end{verbatim}

To change characters of Chinese numerals for each digit, set \kansujichar parameter:

\begin{verbatim}
\ltjsetparameter{kansujichar={1,`壹}}
\ltjsetparameter{kansujichar={7,`漆}}
\ltjsetparameter{kansujichar={0,`零}}
\kansuji1701
\end{verbatim}

9.2 \inhibitglue, \disinhibitglue

\inhibitglue suppresses the insertion of a glue/kern soecified in JFM at the place. The following is an example, using a special JFM that there will be a glue between the beginning of a box and "あ", and also between "あ" and "ウ".

\begin{verbatim}
\jfont=g=HaranoAjiMincho-Regular:jfm=test \g
\bbox{\hbox{ああ\inhibitglue ウ}}
\inhibitglue\par\noindent あ1
\par\inhibitglue\noindent あ2
\par\noindent\inhibitglue あ3
\par 4) \inhibitglue 5
\par\rule{\noindent あoff\inhibitglue ice
\end{verbatim}

With the help of this example, we remark the specification of \inhibitglue:

- The call of \inhibitglue in the (internal) vertical mode is simply ignored.
- \inhibitglue does not suppress \kanjiskip or \xkanjiskip.
• The call of `\inhibitglue` in the (restricted) horizontal mode is only effective on the spot; does not get over boundary of paragraphs. Moreover, `\inhibitglue` cancels ligatures and kernings, as shown in the last line of above example.

• The call of `\inhibitglue` in math mode is just ignored.

`\disinhibitglue` suppresses the effect of `\inhibitglue`. In other words, `\disinhibitglue` allows the insertion of a glue/kern specified by JFM. If `\inhibitglue` and `\disinhibitglue` both specified at the same time, the latest one is effective. This command is added in the version 20201224.0.

Note that `\disinhibitglue` also cancels ligatures and kernings.

9.3 `\ltjfakeboxbdd, \ltjfakeparbegin`

Sometimes 'parbdd' and 'boxbdd' specifications look like “fail”, especially in paragraphs inside list environments. This is because \everypar inserts some nodes such as boxes and kerns, so the “first letter” in a paragraph is in fact not the first letter.

```
\parindent1zw
\noindent ああああああああ\par % for comparison ああああああああ

\everypar{null}
\ltjfakeparbegin

\everypar{null} \ltjfakeparbegin

\everypar{null}
```

`\ltjfakeboxbdd` and `\ltjfakeparbegin` primitives resolve this situation.

• `\ltjfakeparbegin` creates a node which indicates “beginning of an indented paragraph” to the insertion process of `JAglue`.

• `\ltjfakeboxbdd` creates a node which indicates “beginning/ending of a box” to the insertion process of `JAglue`.

As an example, the example above can be improved as follows:

```
\parindent1zw
\noindent ああああああああ\par % for comparison ああああああああ

\everypar{null}\ltjfakeparbegin

\everypar{null}
```

9.4 `\insertxkanjiskip, \insertkanjiskip`

There are some situations which one wants to insert `xkanjiskip` manually. A simple approach is to use `\hskip\ltjgetparameter{xkanjiskip}`, but this approach has several weak points. To cope with these weak points, Lua\TeX-ja defines a command `\insertxkanjiskip` which inserts `xkanjiskip` glue manually, from the version 20201224.0.

• “\insertxkanjiskip” (without any keyword) uses the value of `xkanjiskip` at the place.

• “\insertxkanjiskip late” (with “late” keyword) uses the value of `xkanjiskip` at the end of a paragraph/hbox.

See the example below.

```
\ltjsetparameter{xkanjiskip=0.25zw}
\insertxkanjiskip
あ (\% 0.5zw (from JFM)
\insertxkanjiskip あ (\% 0.25zw (xkanjiskip at here)
\insertxkanjiskip late あ (\% 0.25zw (xkanjiskip at EOP)
\insertxkanjiskip late あ (\% 1.25zw (xkanjiskip at EOP)あ (あ (あ (あ
\ltjsetparameter{xkanjiskip=1.25zw}
\insertxkanjiskip
\insertxkanjiskip
\insertxkanjiskip late あ (\% 1.25zw (xkanjiskip at here)
\insertxkanjiskip late あ (\% 1.25zw (xkanjiskip at EOP)
\insertxkanjiskip late あ (\% At the end of the paragraph (EOP), xkanjiskip is 1.25\zw.
```
There is a similar command `\insertkanjiskip` (kanjiskip instead of vkanjiskip) is also defined. Note that any shorthand form of `\insert[x]kanjiskip` are not defined by LuaTeX-ja.

9.5 `\ltjdeclarealtfont`

Using `\ltjdeclarealtfont`, one can “compose” more than one Japanese fonts. This `\ltjdeclarealtfont` uses in the following form:

```
\ltjdeclarealtfont⟨base_font_cs⟩⟨alt_font_cs⟩{(range)}
```

where `⟨base_font_cs⟩` and `⟨alt_font_cs⟩` are defined by `\jfont`. Its meaning is

If the current Japanese font is `⟨base_font_cs⟩`, characters which belong to `⟨range⟩` is typeset by another Japanese font `⟨alt_font_cs⟩`, instead of `⟨base_font_cs⟩`.

Here `⟨range⟩` is a comma-separated list of character codes, but also accepts negative integers: −n (n ≥ 1) means that all characters of character classes n, with respect to JFM used by `⟨base_font_cs⟩`. Note that characters which do not exist in `⟨alt_font_cs⟩` are ignored.

For example, if `\hoge` uses `jfm-ujis.lua`, the standard JFM of LuaTeX-ja, then

```
\ltjdeclarealtfont\hoge\piyo("3000-"30FF, {-1}-{-1})
```

does

If the current Japanese font is `\hoge`, U+3000–U+30FF and characters in class 1 (ideographic opening brackets) are typeset by `\piyo`.

Note that specifying negative numbers needs specification like `{−1}–{−1}`, because simple “−1” is treated as the range between 0 and 1.

```
\gtfamily\large
\jchar"¥",\ltjchar"¥",\ltjjchar"¥" \% default: ALchar
\jchar"α",\ltjchar"α",\ltjjchar"α" \% default: JAchar
\jchar"g",\ltjchar"g",\ltjjchar"g" \% ALchar unless \ltjjchar
```

10 Commands for \LaTeX\textit{2ε}

10.1 Loading Japanese fonts in \LaTeX\textit{2ε}

From version 20190107, Lua\TeX-ja does not load Japanese fonts for horizontal direction and that for vertical direction at same time, to reduce the number of loaded fonts. This will save time for typesetting and memory consumption of Lua side ([11]).

- `\selectfont` loads (and chooses) only the Japanese font for the current direction, and does not load the Japanese font for other direction (Lua\TeX-ja only detects its size and JFM, to calculate the amount of shifting the baseline).

- Direction changing commands (\yoko, \tate, \dtou, \utod) are patched to include the following process:

  If the Japanese font for new direction is not loaded, Lua\TeX-ja loads it automatically.

Original commands are saved as `\ltj@@orig@yoko` etc.

- Specifying Japanese font command which is defined by `\jfont`, `\tfont`, or `\DeclareFixedFont` directly actually loads (and selects) the Japanese font. For example, `JAchars` in `\box0` will be typeset in `\HOGE`, in the following code:

```
% in horizontal direction (\yoko)
\DeclareFixedFont\HOGE{JT3}{gt}{m}{n}{12} \% JT3: for vertical direction
\HOGE
\setbox0=\hbox{\tate あいう}
```
10.2 Patch for NFSS2

Japanese patch for NFSS2 in LuaTEX-ja is based on p1fonts.dtx which plays the same role in p\TeX\ 2e. We will describe commands which are not described in Subsection 3.1.

Additional dimensions

Like \p\TeX\ 2e, LuaTEX-ja defines the following dimensions for information of current Japanese font:

\chi (height), \cd (depth), \cH (sum of former two),
\cw (width), \cvs (lineskip), \chs (equals to \cw)

and its \\normalsize version:

\Ch (height), \Cd (depth), \Cw (width),
\Cvs (equals to \baselineskip), \Chs (equals to \Cw).

Note that \cw and \cH may differ from \zw and \zh respectively. On the one hand the former dimensions are determined from a character whose character class is zero, but on the other hand \zw and \zh are specified by JFM.

```
\DeclareYokoKanjiEncoding{⟨encoding⟩}{⟨text-settings⟩}{⟨math-settings⟩}
\DeclareTateKanjiEncoding{⟨encoding⟩}{⟨text-settings⟩}{⟨math-settings⟩}
```

In NFSS2 under LuaTEX-ja, distinction between alphabetic fonts and Japanese fonts are only made by their encodings. For example, encodings OT1 and T1 are encodings for alphabetic fonts, and Japanese fonts cannot have these encodings. These command define a new encoding scheme for Japanese font families.

```
\DeclareKanjiEncodingDefaults{⟨text-settings⟩}{⟨math-settings⟩}
\DeclareKanjiSubstitution{⟨encoding⟩}{⟨family⟩}{⟨series⟩}{⟨shape⟩}
\DeclareErrorKanjiFont{⟨encoding⟩}{⟨family⟩}{⟨series⟩}{⟨shape⟩}{⟨size⟩}
```

The above 3 commands are just the counterparts for \DeclareFontEncodingDefaults and others.

```
\reDeclareMathAlphabet{⟨unified-cmd⟩}{⟨al-cmd⟩}{⟨ja-cmd⟩}
\DeclareRelationFont{⟨ja-encoding⟩}{⟨ja-family⟩}{⟨ja-series⟩}{⟨ja-shape⟩}{⟨al-encoding⟩}{⟨al-family⟩}{⟨al-series⟩}{⟨al-shape⟩}
```

This command sets the “accompanied” alphabetic font (given by the latter 4 arguments) with respect to a Japanese font given by the former 4 arguments.

```
\SetRelationFont
\userelfont
```

(Only) at the next call of \selectfont, change current alphabetic font encoding/family/... to the ‘accompanied’ alphabetic font family with respect to current Japanese font family, which was set by \DeclareRelationFont or \SetRelationFont.

The following is an example of \SetRelationFont and \userelfont:

```latex
\maketitle
\SetRelationFont{JY3}{\k@family}{m}{n}{TU}{lmss}{m}{n}
\% \k@family: current Japanese font family
\userelfont\selectfont あいう abc
```

```
\adjustbaseline
```

In \p\TeX\ 2e, \adjustbaseline sets \baselineshift to match the vertical center of "M" and that of "漢" in vertical typesetting:
\[
\text{\baselineshift} \leftarrow \frac{(h_M + d_M) - (h_\text{\normalfont{漢}} + d_\text{\normalfont{漢}})}{2} + d_\text{\normalfont{漢}} - d_M,
\]

where $h_a$ and $d_a$ denote the height of "a" and the depth, respectively. In LuaTEX-ja, this \adjustbaseline does similar task, namely setting the \baselineshift parameter (a Japanese character whose character class is zero is used, instead of "漢").
日本国民は、正当に選挙された国会における代表者を通じて行動し、……

日本国民は、正当に選挙された国会における代表者を通じて行動し、……

Figure 8. An example of DeclareAlternateKanjiFont

\fontfamily{⟨family⟩}
As in \LaTeX{} 2ε, this command changes current font family (alphabetic, Japanese, or both) to ⟨family⟩. See Subsection 10.3 for detail.

\fontshape{⟨shape⟩}, \fontshapeforce{⟨shape⟩}
As in \LaTeX{} 2ε, this command changes current alphabetic font shape according to shape change rules. Traditionally, \fontshape changes also current Japanese font shape always. However, this leads a lot of \LaTeX{} font warning like

\begin{verbatim}
Font shape "JY3/mc/m/it" undefined
using "JY3/mc/m/n" instead on ....
\end{verbatim}

when \itshape is called, because almost all Japanese fonts only have shape "n", and \itshape calls \fontshape.

Lua\TeX{}-ja 20200323.0 change the behavior. Namely, \fontshape{⟨shape⟩} and \fontshapeforce{⟨shape⟩} change current Japanese font shape, only if the required shape (according to shape changing rules) or ⟨shape⟩ is available in current Japanese font family/series. When this is not the case, an info such as

\begin{verbatim}
Kanji font shape JY3/mc/m/it' undefined
No change on ....
\end{verbatim}

is issued instead of a warning.

\kanjishape{⟨shape⟩}, \kanjishapeforce{⟨shape⟩}
\kanjishape{⟨shape⟩} changes current Japanese font shape according to shape change rules, and \kanjishapeforce{⟨shape⟩} changes current Japanese font shape to ⟨shape⟩, regardless of the rules. Hence \kanjishape{it} produces a warning

\begin{verbatim}
Font shape "JY3/mc/m/it" undefined
using "JY3/mc/m/n" instead on ....
\end{verbatim}

which is not produced by \fontshape{it}.

\DeclareAlternateKanjiFont{⟨base-encoding⟩}{⟨base-family⟩}{⟨base-series⟩}{⟨base-shape⟩}
{⟨alt-encoding⟩}{⟨alt-family⟩}{⟨alt-series⟩}{⟨alt-shape⟩}{⟨range⟩}
As \ltjdeclarealtfont (Subsection 9.5), characters in ⟨range⟩ of the Japanese font (we say the base font) which specified by first 4 arguments are typeset by the Japanese font which specified by fifth to eighth arguments (we say the alternate font). An example is shown in Figure 8.

• In \ltjdeclarealtfont, the base font and the alternate font must be already defined. But this \DeclareAlternateKanjiFont is not so. In other words, \DeclareAlternateKanjiFont is effective only after current Japanese font is changed, or only after \selectfont is executed.

• ...

40
Furthermore, LuaTeX-ja applies patches which enables NFSS2 commands, such as \DeclareSymbolFont and \SetSymbolFont, to specify Japanese fonts as math fonts.

Specifying disablejfam option in \usepackage prevents applying these patches. Hence one cannot write Japanese Characters in math mode directly if disablejfam option is specified. The code below does not work either:
\begin{verbatim}
\DeclareSymbolFont{mincho}{JY3}{mc}{m}{n}
\DeclareSymbolFontAlphabet{\mathmc}{mincho}
\end{verbatim}

10.3 Detail of \fontfamily command

In this subsection, we describe when \fontfamily{\langle family \rangle} changes current Japanese/alphabetic font family. Basically, current Japanese font family is changed to \langle family \rangle if it is recognized as a Japanese font family, and similar with alphabetic font family. There is a case that current Japanese/alphabetic font family are both changed to \langle family \rangle, and another case that \langle family \rangle isn’t recognized as a Japanese/alphabetic font family either.

■ Recognition as Japanese font family

First, whether Japanese font family will be changed is determined in following order. This order is very similar to \fontfamily in p\TeX\ 2\kappa, but we re-implemented in Lua. We use an auxiliary list \text{\textit{N}}_j.

1. If the family \langle family \rangle has been defined already by \DeclareKanjiFamily, \langle family \rangle is recognized as a Japanese font family. Note that \langle family \rangle need not be defined under current Japanese font encoding.

2. If the family \langle family \rangle has been listed in a list \text{\textit{N}}_j, this means that \langle family \rangle is not a Japanese font family.

3. If the luatexja-fontspec package is loaded, we stop here, and \langle family \rangle is not recognized as a Japanese font family.

If the luatexja-fontspec package is not loaded, now LuaTeX-ja looks whether there exists a Japanese font encoding \langle enc \rangle such that a font definition named \langle enc \rangle\langle family \rangle.fd (the file name is all lowercase) exists. If so, \langle family \rangle is recognized as a Japanese font family (the font definition file won’t be loaded here). If not, \langle family \rangle is not a Japanese font family, and \langle family \rangle is appended to the list \text{\textit{N}}_j.

■ Recognition as alphabetic font family

Next, whether alphabetic font family will be changed is determined in following order. We use auxiliary lists \text{\textit{F}}_A and \text{\textit{N}}_A.

1. If the family \langle family \rangle has been listed in a list \text{\textit{F}}_A, \langle family \rangle is recognized as an alphabetic font family.

2. If the family \langle family \rangle has been listed in a list \text{\textit{N}}_A, this means that \langle family \rangle is not an alphabetic font family.

3. If there exists an alphabetic font encoding such that the family \langle family \rangle has been defined under it, \langle family \rangle is recognized as an alphabetic font family, and to memorize this, \langle family \rangle is appended to the list \text{\textit{F}}_A.

4. Now LuaTeX-ja looks whether there exists an alphabetic font encoding \langle enc \rangle such that a font definition named \langle enc \rangle\langle family \rangle.fd (the file name is all lowercase) exists. If so, current alphabetic font family will be changed to \langle family \rangle (the font definition file won’t be loaded here). If not, current alphabetic font family won’t be changed, and \langle family \rangle is appended to the list \text{\textit{N}}_A.

Also, each call of \DeclareFontFamily after loading of LuaTeX-ja makes the second argument (family) is appended to the list \text{\textit{F}}_A.

The above order is very similar to \fontfamily in p\TeX\ 2\kappa, but more complicated (clause 3.). This is because p\TeX\ 2\kappa is a format however LuaTeX-ja is not, hence LuaTeX-ja does not know calls of \DeclareFontFamily before itself is loaded.
Table 15. strut

<table>
<thead>
<tr>
<th>box</th>
<th>direction</th>
<th>width</th>
<th>height</th>
<th>depth</th>
<th>user command</th>
</tr>
</thead>
<tbody>
<tr>
<td>\ystrutbox</td>
<td>yoko</td>
<td>0</td>
<td>0.7\baselineskip</td>
<td>0.3\baselineskip</td>
<td>\ystrut</td>
</tr>
<tr>
<td>\tstrutbox</td>
<td>tate, utod</td>
<td>0</td>
<td>0.5\baselineskip</td>
<td>0.5\baselineskip</td>
<td>\tstrut</td>
</tr>
<tr>
<td>\dstrutbox</td>
<td>dtou</td>
<td>0</td>
<td>0.7\baselineskip</td>
<td>0.3\baselineskip</td>
<td>\dstrut</td>
</tr>
<tr>
<td>\zstrutbox</td>
<td>—</td>
<td>0</td>
<td>0.7\baselineskip</td>
<td>0.3\baselineskip</td>
<td>\zstrut</td>
</tr>
</tbody>
</table>

Remarks

Of course, there is a case that ⟨family⟩ is not recognized as a Japanese font family, nor an alphabetic font family. In this case, LuaTEX-ja treats “the argument ⟨family⟩ is wrong”, so set both current alphabetic and Japanese font family to ⟨family⟩, to use the default family for font substitution.

10.4 Notes on \DeclareTextSymbol

From \TeX\ 2017/01/01, the standard encoding of \LaTeX\ is changed to the TU encoding. This means that symbols defined by T1 and TS1 encodings can be used without loading any package. To produces these symbols in alphabetic fonts in LuaTEX-ja, LuaTEX-ja patches \DeclareTextSymbol, and reloads tuenc.def.

Under original definition of \DeclareTextSymbol, internal commands which is defined by \DeclareTextSymbol (such as \T1\textquotedblleft) are chardef tokens. However, this no longer holds in LuaTEX-ja; for example, the meaning of \T1\textquotedblleft is \ltjalchar8220␣.

10.5 \strutbox

As \TeX\ (2017/04/08 or later), \strutbox is a macro which is expanded to one of \ystrutbox, \tstrutbox, and \dstrutbox (all of them are shown in Table 15), according to the current direction. Similarly, \strut now uses one of these boxes.

11 expl3 interface

This section describes expl3 interfaces provided by \LaTeX-ja. All of them belong to te platex module, since they are provided for compatibility with Japanese \TeX. Note that commands which are marked with dagger (“†”) are additions by LuaTEX-ja.

\LaTeX\direction{yoko}: \LaTeX\direction{tate}: \LaTeX\direction{dtou}:

Synonyms for \yoko, \tate and \dtou, respectively.

\LaTeX\if\direction{yoko}{}\{true code\}\{false code\}
Tests if the current direction is \yoko (horizontal writing).

\LaTeX\if\direction{tate}{nomath}{}\{true code\}\{false code\}
Tests if the current direction is \tate (vertical writing).

\LaTeX\if\direction{tate}{math}{}\{true code\}\{false code\}
Tests if the current direction is \dtou.

\LaTeX\if\direction{tate}\{true code\}\{false code\}
Tests if the current direction is \tate or \dtou.

\LaTeX\if\direction{dtou}\{true code\}\{false code\}
Tests if the current direction is \dtou.
12 Addon packages

Lua\TeX\-ja has several addon packages. These addons are written as \LaTeX\ packages, but luatexja-otf and luatexja-adjust can be loaded in plain \LaTeX\ by \texttt{\input}.

12.1 \texttt{luatexja-fontspec}

As described in Subsection 3.2, this optional package provides the counterparts for several commands defined in the fontspec package (requires fontspec v2.4). In addition to OpenType font features in the original fontspec, the following "font features" specifications are allowed for the commands of Japanese version:

\begin{verbatim}
\texttt{CID=⟨name⟩, JFM=⟨name⟩, JFM-var=⟨name⟩}
\end{verbatim}

These 3 keys correspond to \texttt{cid}, \texttt{jfm} and \texttt{jfmvar} keys for \texttt{\jfont} and \texttt{\tfont} respectively. See Subsections 7.1 and 7.4 for details of \texttt{cid}, \texttt{jfm} and \texttt{jfmvar} keys.

The \texttt{CID} key is effective only when with \texttt{NoEmbed} described below. The same \texttt{JFM} cannot be used in both horizontal Japanese fonts and vertical Japanese fonts, hence the \texttt{JFM} key will be actually used in \texttt{YokoFeatures} and \texttt{TateFeatures} keys.

\begin{verbatim}
\texttt{NoEmbed}
\end{verbatim}

By specifying this key, one can use "name-only" Japanese font which will not be embedded in the output PDF file. See Subsection 7.4.

\begin{verbatim}
\texttt{Kanjiskip=⟨bool⟩}
\end{verbatim}
日本国民は、正当に選挙された国会における代表者を通じて行動し、われらとわれらの子孫のために、
諸国民との協和による成果と、わが国全土にわたって自由のもたらす恵沢を確保し、……

Figure 10. An example of AltFont

TateFeatures={⟨features⟩}, TateFont=⟨font⟩
The TateFeatures key specifies font features which are only turned on in vertical writing, such as
Style=VerticalKana (vkna feature). Similarly, the TateFont key specifies the Japanese font which
will be used only in vertical writing. A demonstrarion is shown in Figure 9.

YokoFeatures={⟨features⟩}
The YokoFeatures key specifies font features which are only turned on in horizontal writing. A
demonstrarion is shown in Figure 9.

AltFont
As \ltjdeclarealtfont (Subsection 9.5) and \DeclareAlternateKanjiFont (Subsection 10.2), with
this key, one can typeset some Japanese characters by a different font and/or using different features.
The AltFont feature takes a comma-separated list of comma-separated lists, as the following:
AltFont = { ...
{ Range=⟨range⟩, ⟨features⟩ },
{ Range=⟨range⟩, Font=⟨font name⟩, ⟨features⟩ },
{ Range=⟨range⟩, Font=⟨font name⟩ },
...
} Each sublist should have the Range key (sublist which does not contain Range key is simply ignored).
A demonstrarion is shown in Figure 10.

■Remark on AltFont, YokoFeatures, TateFeatures keys
In AltFont, YokoFeatures, TateFeatures keys, one cannot specify per-shape settings such as
BoldFeatures. For example,
AltFont = { ...
{ Font=HogeraMin-Light, BoldFont=HogeraMin-Bold,
  Range="3000-"30FF, BoldFeatures={Color=FF1900} } }
does not work. Instead, one have to write
UprightFeatures = { 
  AltFont = { { Font=HogeraMin-Light, Range="3000-"30FF, } },
  BoldFeatures = { 
  AltFont = { { Font=HogeraMin-Bold, Range="3000-"30FF, Color=FF1900 } },
  }
}

On the other hand, YokoFeatures, TateFeatures and TateFont keys can be specified in each list in
the AltFont key. Also, one can specify AltFont inside YokoFeatures, TateFeatures.
The value of \kanjiskip is 0 pt $+\tfrac{1}{5}$ em in this figure, for making the difference obvious.

Figure 11. Line adjustment

Note that features which are specified in \YokoFeatures and \TateFeatures are always interpreted after other "direction-independent" features. This explains why \addjfontfeatures at line 6 in Figure 9 has no effect, because a color specification is already done in \YokoFeatures and \TateFeatures keys.

12.2 \texttt{luatexja-otf}

This optional package supports typesetting glyphs by specifying a CID number. The package \texttt{luatexja-otf} offers the following 2 low-level commands:

\begin{verbatim}
\CID{(number)}
\end{verbatim}

Typeset a glyph whose CID number is \texttt{(number)}. If the Japanese font is neither Adobe-Japan1, Adobe-GB1, Adobe-CNS1, Adobe-Korea1, nor Adobe-KR CID-keyed font, \LaTeX-ja treats that \texttt{(number)} is a CID number of Adobe-Japan1 character collection, and tries to typeset a "most suitable glyph".

Note that if the Japanese font is loaded using the HarfBuzz library, this \texttt{\CID} command does not work.

\begin{verbatim}
\UTF{(hex_number)}
\end{verbatim}

Typeset a character whose character code is \texttt{(hex_number)} (in hexadecimal). This command is similar to \texttt{\char\{(hex_number\)}}, but please remind remarks below.

This package automatically loads \texttt{luatexja-ajmacros.sty}, which is slightly modified version of \texttt{ajmacros.sty}\textsuperscript{11}. Hence one can use macros which are defined in \texttt{ajmacros.sty}, such as \texttt{\aj半角}.

\begin{itemize}
\item \textbf{Remarks} Characters by \texttt{\CID} and \texttt{\UTF} commands are different from ordinary characters in the following points:
\begin{itemize}
\item Always treated as \texttt{JAchars}.
\item In vertical direction, \texttt{\vrt/vert2} feature are automatically applied to characters by \texttt{\UTF}, regardless these feature are not activated in current Japanese font.
\item Processes for supporting other OpenType features (for example, glyph replacement and kerning) by the \texttt{luaotfload} package is not performed to these characters.
\end{itemize}
\end{itemize}

\begin{itemize}
\item \textbf{Additional syntax of JFM} The package \texttt{luatexja-otf} extends the syntax of JFM; the entries of \texttt{chars} table in JFM now allows a string in the form \texttt{'AJ1-xxx'}, which stands for the character whose CID number in Adobe-Japan1 is \texttt{xxx}.

This extended notation is used in the standard JFM \texttt{jfm-ujis.1ua} to typeset halfwidth Hiragana glyphs (CID 516–598) in halfwidth.
\end{itemize}

12.3 \texttt{luatexja-adjust}

(see Japanese version of this manual)

\textsuperscript{11}Useful macros by INOUE Koich!, for the \texttt{japanese-otf} package.
12.4 luatexja-ruby

This addon package provides functionality of “ruby” (furigana) annotations using callbacks of Lua\TeX-ja. There is no detailed manual of luatexja-ruby.sty in English. (Japanese manual is another PDF file, luatexja-ruby.pdf.)

**Group-ruby**  By default, ruby characters (the second argument of \ruby) are attached to base characters (the first argument), as one object. This type of ruby is called group-ruby.

```
東西線ruby(妙典){みようでん}駅は……
東西線のruby(妙典){みようでん}駅は……
東西線的ruby(妙典){みようでん}という駅……
東西線ruby(葛西){かさい}駅は……
```

As the above example, ruby hangover is allowed on the Hiragana before/after its base characters.

**Mono-ruby**  To attach ruby characters to each base characters (mono-ruby), one should use \ruby multiple times:

```
東西線のruby(妙){みよう}{典}{でん}駅は……
東西線のみよう妙でん典駅は……
東西線のみようでん妙典という駅……
東西線的葛西{かさい}駅は……
```

**Jukugo-ruby**  Vertical bar | denotes a boundary of groups.

```
ruby(妙){みよう}{典}{でん}
ruby(葛){か}{さい}
```

If there are multiple groups in one \ruby call, A linebreak between two groups is allowed.

```
vbox{hspace=6zw}\noindent
 hbox to 2.5zw}{ruby(京急}{蒲}{田}{けいきゆう}{かま}{た}
 hbox to 2.5zw}{ruby(京急}{蒲}{田}{けいきゆう}{かま}{た}
 hbox to 3zw}{ruby(京急}{蒲}{田}{けいきゆう}{かま}{た}
```

If the width of ruby characters are longer than that of base characters, \ruby automatically selects the appropriate form among the line-head form, the line-middle form, and the line-end form.

```
vbox{hspace=8zw}\noindent
 null\kern3zw ……\noindent\rubysize=8zw\noindent
 null\kern3zw …… \noindent\rubysize=8zw\noindent
```

12.5 lltjext.sty

p\TeX supplies additional macros for vertical writing in the plext package. The lltjext package which we want to describe here is the Lua\TeX-ja counterpart of the plext package.

**tabular, array, minipage environments**  These environments are extended by <dir>, which specifies the direction, as follows:

```
\begin{tabular}<dir>\[pos]\{table spec\} ... \end{tabular}
\begin{array}<dir>\[pos]\{table spec\} ... \end{array}
\begin{minipage}<dir>\[pos]\{width\} ... \end{minipage}
```

This option permits one of the following five values. If none of them is specified, the direction inside the environment is same as that outside the environment.

- y yoko direction (horizontal writing)
\parbox<dir>[\width]{\content}
\parbox command is also extended by <\dir>.

\pbox<dir>[(\width)][\pos]{\content}
This commands typeset \content in LR-mode, in \dir direction. If \width is positive, the width of the box becomes this \width. In this case, \content will be aligned to left (when \pos is 1), center (c), or right (r).

picture environment
picture environment also extended by <\dir>, as follows:
\begin{picture}<dir>(x\_size, y\_size)(x\_offset, y\_offset)
...\end{picture}
\rensujij{(\pos)}{(\content)}, \rensujiskip
\Kanji{(counter\_name)}
\kasen{(\content)}, \bou{(\content)}, \boutenchar

参照番号

12.6 luatexja-preset

As described in Subsection 3.3, One can load the luatexja-preset package to use several “presets” of Japanese fonts. This package provides functions in a part of japanese-otf package (changing fonts) and a part of PXchfon package (presets) by Takayuki Yato.

Options which are given in \usepackage but not described in this subsection are simply passed to the luatexja-fontspec. For example, the line 5 in below example is equivalent to lines 1–3.
\usepackage[no-math][fontspec]
\usepackage[match]{luatexja-fontspec}
\usepackage[kozuka-pr\'on]{luatexja-preset}
\usepackage[no-math,match,kozuka-pr\'on]{luatexja-preset}

12.6.1 General Options

fontspec (enabled by default)
With this option, Japanese fonts are selected using functionality of the luatexja-fontspec package. This means that the fontspec package is automatically loaded by this package.

If you need to pass some options to fontspec, you can load fontspec manually before luatexja-preset:
\usepackage[no-math]{fontspec}
\usepackage[...]{luatexja-preset}

\textsuperscript{12}if nfssonly option is not specified; in this case these options are simply ignored.
With this option, selecting Japanese fonts won’t be performed using the functionality of the `fontspec` package, but only standard NFSS2 (hence without \addfontfeatures etc.). This option is ignored when `luatexja-fontspec` package is loaded.

When this option is specified, `fontspec` and `luatexja-fontspec` are not loaded by default. Nevertheless, the package `fontspec` can coexist with the option, as the following:

\usepackage[fontspec]
\usepackage[hiragino-pron,nfssonly]{luatexja-preset}

In this case, one can use \setmainfont etc. to select alphabetic fonts.

If this option is specified, usual family-changing commands such as \rmfamily, \textttfamily, \textitfamily, ... also change Japanese font family. This option is passed to `luatexja-fontspec`, if `fontspec` option is specified.

The negation of `deluxe` option. Use one-weighted mincho and gothic font families. This means that \mcfamily\bfseries, \gtfamily\bfseries and \gtfamily\mdseries use the same font.

Use the mincho family with three weights (light, medium, and bold), the gothic family with three weights (medium, bold, and extra bold), and `rounded gothic`\footnote{Provided by \mgfamily and \textmg, because “rounded gothic” is called maru gothic (丸ゴシック) in Japanese.}. Mincho light and gothic extra bold can be by \mcfamily\ltseries and \gtfamily\ebseries, respectively.

- Some presets do not have the light weight of mincho. In this case, we substitute the medium weight for the light weight.
- `luatexja-preset` does not produce an error (only produces a warning), even if (one of) fonts for \mcfamily\ltseries, \gtfamily\ebseries, \mgfamily do not exist.

Use horizontal/vertical kana alternates, and define a command \rubyfamily to use kana characters designed for ruby.

Substitute bold series of gothic for bold series of mincho. If `nodeluxe` option is enabled, medium series of gothic is also changed, since we use same font for both series of gothic.

\texttt{jis90}, \texttt{90jis}
Use JIS X 0208:1990 glyph variants if possible.

\texttt{jis2004}, \texttt{2004jis}
Use JIS X 0213:2004 glyph variants if possible.

\texttt{jfm\_yoko=(jfm)}
Use the JFM jfm-(jfm).lua for horizontal direction, instead of jfm-ujis.1ua (default JFM).

\texttt{jfm\_tate=(jfm)}
Use the JFM jfm-(jfm).lua for vertical direction, instead of jfm-ujisv.1ua (default JFM).

\texttt{jis}
Same as jfm\_yoko=jis.

Note that jis\texttt{90}, jis\texttt{90jis}, jis\texttt{2004} and jis\texttt{2004jis} only affect with mincho, gothic (and, possibly rounded gothic) families defined by this package. We didn’t taken account of when more than one options among them are specified.
### 12.6.2 Presets which support multi weights

Besides bizud, haranoaji, morisawa-pro, and morisawa-pr6n presets, fonts are specified by font name, not by file name. In following tables, starred fonts (e.g. KozGo...-Regular) are used for medium series of *gothic*, *if and only if* *deluxe* option is specified.

**Kozuka Pro/Pr6N fonts are bundled with Adobe’s software, such as Adobe InDesign. There is not rounded gothic family in Kozuka fonts.**

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>kozuka-pro</th>
<th>kozuka-pr6</th>
<th>kozuka-pr6n</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>mincho</em></td>
<td>light</td>
<td>KozMinPro-Light</td>
<td>KozMinProVI-Light</td>
<td>KozMinPr6N-Light</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>KozMinPro-Regular</td>
<td>KozMinProVI-Regular</td>
<td>KozMinPr6N-Regular</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>KozMinPro-Bold</td>
<td>KozMinProVI-Bold</td>
<td>KozMinPr6N-Bold</td>
</tr>
<tr>
<td><em>gothic</em></td>
<td>medium</td>
<td>KozGoPro-Regular*</td>
<td>KozGoProVI-Regular*</td>
<td>KozGoPr6N-Regular*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KozGoPro-Medium</td>
<td>KozGoProVI-Medium</td>
<td>KozGoPr6N-Medium</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>KozGoPro-Bold</td>
<td>KozGoProVI-Bold</td>
<td>KozGoPr6N-Bold</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>KozGoPro-Heavy</td>
<td>KozGoProVI-Heavy</td>
<td>KozGoPr6N-Heavy</td>
</tr>
<tr>
<td><em>rounded gothic</em></td>
<td></td>
<td>KozGoPro-Heavy</td>
<td>KozGoProVI-Heavy</td>
<td>KozGoPr6N-Heavy</td>
</tr>
</tbody>
</table>

**Hiragino fonts (except Hiragino Mincho W2) are bundled with Mac OS X 10.5 or later. Note that fonts for gothic extra bold (HiraKakuStd[N]-W8) only contains characters in Adobe-Japan1-3 character collection, while others contains those in Adobe-Japan1-5 character collection.**

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>hiragino-pro</th>
<th>hiragino-pron</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>mincho</em></td>
<td>light</td>
<td>Hiragino Mincho Pro W2</td>
<td>Hiragino Mincho ProN W2</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>Hiragino Mincho Pro W3</td>
<td>Hiragino Mincho ProN W3</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Hiragino Mincho Pro W6</td>
<td>Hiragino Mincho ProN W6</td>
</tr>
<tr>
<td><em>gothic</em></td>
<td>medium</td>
<td>Hiragino Kaku Gothic Pro W3</td>
<td>Hiragino Kaku Gothic ProN W3*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hiragino Kaku Gothic Pro W6</td>
<td>Hiragino Kaku Gothic ProN W6</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Hiragino Kaku Gothic Pro W6</td>
<td>Hiragino Kaku Gothic ProN W6</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>Hiragino Kaku Gothic Std W8</td>
<td>Hiragino Kaku Gothic StdN W8</td>
</tr>
<tr>
<td><em>rounded gothic</em></td>
<td></td>
<td>Hiragino Maru Gothic Pro W4</td>
<td>Hiragino Maru Gothic ProN W4</td>
</tr>
</tbody>
</table>

**BIZ UD fonts (by Morisawa Inc.) bundled with Windows 10 October 2018 Update.**

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>mincho</em></td>
<td>BIZ-UDMinchoM.ttc</td>
</tr>
<tr>
<td><em>gothic</em></td>
<td>BIZ-UDGothicR.ttc</td>
</tr>
<tr>
<td>bold</td>
<td>BIZ-UDGothicB.ttc</td>
</tr>
<tr>
<td>extra bold</td>
<td>BIZ-UDGothicB.ttc</td>
</tr>
<tr>
<td><em>rounded gothic</em></td>
<td>BIZ-UDGothicB.ttc</td>
</tr>
</tbody>
</table>
**morisawa-pro** Morisawa Pro (Adobe-Japan1-4) fonts.
**morisawa-pr6n** Morisawa Pr6N (Adobe-Japan1-6, JIS04-savvy) fonts.

<table>
<thead>
<tr>
<th>family series</th>
<th>morisawa-pro</th>
<th>morisawa-pr6n</th>
</tr>
</thead>
<tbody>
<tr>
<td>minche medium</td>
<td>A-OTF-RyuminPro-Light.otf</td>
<td>A-OTF-RyuminPr6N-Light.otf</td>
</tr>
<tr>
<td>bold</td>
<td>A-OTF-FutoMinA101Pro-Bold.otf</td>
<td>A-OTF-FutoMinA101Pr6N-Bold.otf</td>
</tr>
<tr>
<td>gothic medium</td>
<td>A-OTF-GothicBBBPro-Medium.otf</td>
<td>A-OTF-GothicBBBPr6N-Medium.otf</td>
</tr>
<tr>
<td>bold</td>
<td>A-OTF-FutoGoB101Pro-Bold.otf</td>
<td>A-OTF-FutoGoB101Pr6N-Bold.otf</td>
</tr>
<tr>
<td>extra bold</td>
<td>A-OTF-MidashiGoPro-MB31.otf</td>
<td>A-OTF-MidashiGoPr6N-MB31.otf</td>
</tr>
</tbody>
</table>

**yu-win** Yu fonts bundled with Windows 8.1.
**yu-win10** Yu fonts bundled with Windows 10.
**yu-osx** Yu fonts bundled with OSX Mavericks.

<table>
<thead>
<tr>
<th>family series</th>
<th>yu-win</th>
<th>yu-win10</th>
<th>yu-osx</th>
</tr>
</thead>
<tbody>
<tr>
<td>minche light</td>
<td>YuMincho-Light</td>
<td>YuMincho-Light</td>
<td>(YuMincho Medium)</td>
</tr>
<tr>
<td>medium</td>
<td>YuMincho-Regular</td>
<td>YuMincho-Regular</td>
<td>YuMincho Medium</td>
</tr>
<tr>
<td>bold</td>
<td>YuMincho-DemiBold</td>
<td>YuMincho-DemiBold</td>
<td>YuMincho DemiBold</td>
</tr>
<tr>
<td>gothic medium</td>
<td>YuGothic-Regular*</td>
<td>YuGothic-Regular*</td>
<td>YuGothic Medium*</td>
</tr>
<tr>
<td>bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic Bold</td>
</tr>
<tr>
<td>extra bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic Bold</td>
</tr>
<tr>
<td>rounded gothic</td>
<td>YuGothic-Bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic Bold</td>
</tr>
</tbody>
</table>

**moga-mobo** MogaMincho, MogaGothic, and MoboGothic.
**moga-mobo-ex** MogaExMincho, MogaExGothic, and MoboExGothic.

These fonts can be downloaded from [http://yozvox.web.fc2.com/](http://yozvox.web.fc2.com/).

<table>
<thead>
<tr>
<th>family series</th>
<th>default, 90jis option</th>
<th>jis2004 option</th>
</tr>
</thead>
<tbody>
<tr>
<td>minche medium</td>
<td>Moga90Mincho</td>
<td>MogaMincho</td>
</tr>
<tr>
<td>bold</td>
<td>Moga90Mincho Bold</td>
<td>MogaMincho Bold</td>
</tr>
<tr>
<td>gothic medium</td>
<td>Moga90Gothic</td>
<td>MogaGothic</td>
</tr>
<tr>
<td>bold</td>
<td>Moga90Gothic Bold</td>
<td>MogaGothic Bold</td>
</tr>
<tr>
<td>extra bold</td>
<td>Moga90Gothic Bold</td>
<td>MogaGothic Bold</td>
</tr>
<tr>
<td>rounded gothic</td>
<td>Mobo90Gothic</td>
<td>MoboGothic</td>
</tr>
</tbody>
</table>

When **moga-mobo-ex** is specified, the font “MogaEx90Mincho” etc. are used.

**ume** Ume Mincho and Ume Gothic.


<table>
<thead>
<tr>
<th>family series</th>
<th>default</th>
</tr>
</thead>
<tbody>
<tr>
<td>minche medium</td>
<td>Ume Mincho</td>
</tr>
<tr>
<td>bold</td>
<td>Ume Mincho</td>
</tr>
<tr>
<td>gothic medium</td>
<td>Ume Gothic*</td>
</tr>
<tr>
<td></td>
<td>Ume Gothic O5</td>
</tr>
<tr>
<td>bold</td>
<td>Ume Gothic O5</td>
</tr>
<tr>
<td>extra bold</td>
<td>Ume Gothic O5</td>
</tr>
<tr>
<td>rounded gothic</td>
<td>Ume Gothic O5</td>
</tr>
</tbody>
</table>
sourcehan  Source Han Serif and Source Han Sans fonts (Language-specific OTF or OTC)
sourcehan-jp  Source Han Serif JP and Source Han Sans JP fonts (Region-specific Subset OTF)

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>sourcehan</th>
<th>sourcehan-jp</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>light</td>
<td>Source Han Serif Light</td>
<td>Source Han Serif JP Light</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>Source Han Serif Regular</td>
<td>Source Han Serif JP Regular</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Source Han Serif Bold</td>
<td>Source Han Serif JP Bold</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>Source Han Sans Regular*</td>
<td>Source Han Sans JP Regular*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Source Han Sans Medium*</td>
<td>Source Han Sans JP Medium*</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Source Han Sans Bold</td>
<td>Source Han Sans JP Bold</td>
</tr>
<tr>
<td></td>
<td>extra</td>
<td>Source Han Sans Heavy</td>
<td>Source Han Sans JP Heavy</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>Source Han Sans Medium*</td>
<td>Source Han Sans JP Medium*</td>
</tr>
</tbody>
</table>

noto-otc  Noto Serif CJK and Noto Sans CJK fonts (OTC)
noto-otf  Noto Serif CJK and Noto Sans CJK fonts (Language-specific OTF)

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>noto-otc</th>
<th>noto-otf</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>light</td>
<td>Noto Serif CJK Light</td>
<td>Noto Serif CJK JP Light</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>Noto Serif CJK Regular</td>
<td>Noto Serif CJK JP Regular</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Noto Serif CJK Bold</td>
<td>Noto Serif CJK JP Bold</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>Noto Sans CJK Regular*</td>
<td>Noto Sans CJK JP Regular*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Noto Sans CJK Medium*</td>
<td>Noto Sans CJK JP Medium*</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Noto Sans CJK Bold</td>
<td>Noto Sans CJK JP Bold</td>
</tr>
<tr>
<td></td>
<td>extra</td>
<td>Noto Sans CJK Black</td>
<td>Noto Sans CJK JP Black</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>Noto Sans CJK Medium*</td>
<td>Noto Sans CJK JP Medium*</td>
</tr>
</tbody>
</table>

haranoaji  Harano Aji Fonts.
These fonts can be downloaded from [https://github.com/trueroad/HaranoAjiFonts](https://github.com/trueroad/HaranoAjiFonts). There is not rounded gothic family in Harano Aji Fonts.

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>haranoaji</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>light</td>
<td>HaranoAjiMincho-Light.otf</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>HaranoAjiMincho-Regular.otf</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>HaranoAjiMincho-Bold.otf</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>HaranoAjiGothic-Regular.otf*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HaranoAjiGothic-Medium.otf</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>HaranoAjiGothic-Bold.otf</td>
</tr>
<tr>
<td></td>
<td>extra</td>
<td>HaranoAjiGothic-Heavy.otf</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>HaranoAjiGothic-Medium.otf</td>
</tr>
</tbody>
</table>

12.6.3  Presets which do not support multi weights

Next, we describe settings for using only single weight.

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>ipa</th>
<th>ipaex</th>
<th>ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>Ryumin-Light (non-embedded)</td>
<td>IPA Mincho</td>
<td>IPAex Mincho</td>
<td>MS Mincho</td>
</tr>
<tr>
<td>gothic</td>
<td>GothicBBB-Medium (non-embedded)</td>
<td>IPA Gothic</td>
<td>IPAex Gothic</td>
<td>MS Gothic</td>
</tr>
</tbody>
</table>
### 12.6.4 Presets which use HG fonts

We can use HG fonts bundled with Microsoft Office for realizing multiple weights. In the table below, starred fonts (e.g., IPA Gothic*) are used only if jis2004 or nodeluxe option is specified.

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th>ipa-hg</th>
<th>ipaex-hg</th>
<th>ms-hg</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>medium</td>
<td>IPA Mincho</td>
<td>IPAex Mincho</td>
<td>MS Mincho</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>HG Mincho E</td>
<td>HG Mincho E</td>
<td>HG Mincho E</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>IPA Gothic*</td>
<td>IPAex Gothic*</td>
<td>MS Gothic*</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>HG Gothic M</td>
<td>HG Gothic M</td>
<td>HG Gothic M</td>
</tr>
<tr>
<td>rounded gothic</td>
<td>extra bold</td>
<td>HG Soei Kaku Gothic UB</td>
<td>HG Soei Kaku Gothic UB</td>
<td>HG Soei Kaku Gothic UB</td>
</tr>
</tbody>
</table>

Note that HG Mincho E, HG Gothic E, HG Soei Kaku Gothic UB, and HG Maru Gothic PRO are internally specified by:

- **default** by font name (HGMinchoE, etc.).
- **jis90, 90jis** by file name (hgrme.ttc, hgrge.ttc, hgrsgu.ttc, hgrsmp.ttf).
- **jis2004, 2004jis** by file name (hgrme04.ttc, hgrge04.ttc, hgrsgu04.ttc, hgrsmp04.ttf).

### 12.6.5 Define/Use Custom Presets

From version 20170904.0, one can define new presets using `\ltjnewpreset`, and use them by `\ltjapplypreset`. These two commands can only be used in the preamble.

\[\ltjnewpreset\{\langle name\rangle\}\{\langle specification\rangle\}\]

Define new preset \(\langle name\rangle\). This \(\langle name\rangle\) cannot be same as other presets, options described in Sub-subsection 12.6.1, nor following 11 strings:

\(\langle specification\rangle\) is a comma-separated list which consists of other presets and/or the following keys:

- \(\langle mc-l=\rangle\langle font\rangle\) mincho light
- \(\langle mc-m=\rangle\langle font\rangle\) mincho medium
- \(\langle mc-b=\rangle\langle font\rangle\) mincho bold
- \(\langle mc-bx=\rangle\langle font\rangle\) synonym for \(\langle mc-b=\rangle\langle font\rangle\)
- \(\langle gt-m=\rangle\langle font\rangle\) gothic medium
- \(\langle gt-b=\rangle\langle font\rangle\) gothic bold
- \(\langle gt-bx=\rangle\langle font\rangle\) synonym for \(\langle gt-b=\rangle\langle font\rangle\)
- \(\langle gt-eb=\rangle\langle font\rangle\) gothic extra bold
- \(\langle mg-m=\rangle\langle font\rangle\) rounded gothic

\(\langle mc=\rangle\langle font\rangle\) Same as \(\langle mc-l=\rangle\langle font\rangle, \langle mc-m=\rangle\langle font\rangle, \langle mc-b=\rangle\langle font\rangle\)

\(\langle gt=\rangle\langle font\rangle\) Same as \(\langle gt-m=\rangle\langle font\rangle, \langle gt-b=\rangle\langle font\rangle, \langle gt-eb=\rangle\langle font\rangle\)

If deluxe is not specified at loading the package, only \(\langle mc\rangle\) and \(\langle gt\rangle\) keys (among above 11 keys) have a meaning.

\[\ltjnewpreset*\{\langle name\rangle\}\{\langle specification\rangle\}\]

Almost same as `\ltjnewpreset`. However, if \(\langle name\rangle\) matches a preset which already defined, this command simply overwrite it.
\texttt{\textbackslash ltjapplypreset\{\textit{name}\}}

Set Japanese font families using preset \textit{name}.

Note that \texttt{\textbackslash ltjnewpreset} does not "expand" the definition to define a preset. This means that one can write as the following:

\texttt{\textbackslash ltjnewpreset\{hoge\}\{piyo,mc-b=HiraMinProN-W6\}}
\texttt{\textbackslash ltjnewpreset\{piyo\}\{mg-n=HiraMaruProN-W4\}}
\texttt{\textbackslash ltjapplypreset\{hoge\}}

\textbf{Restrictions} Presets which are defined by \texttt{\textbackslash ltjnewpreset} have following restrictions:

\begin{itemize}
  \item One cannot specify non-embedded fonts (such as Ryumin-Light).
  \item Some presets, such as \texttt{ipa-hg}, have a feature that fonts are changed according to whether \texttt{90jis} or \texttt{jis2004} is specified. This feature is not usable in presets which are defined by \texttt{\textbackslash ltjnewpreset}.
\end{itemize}
Part III
Implementations

13 Storing Parameters

13.1 Used dimensions, attributes and whatsit nodes

Here the following is the list of dimensions and attributes which are used in Lua\TeX-ja.

\jQ (dimension) \jQ is equal to 1 \textbackslash Q = 0.25 \text{ mm}, where "Q" (also called "級") is a unit used in Japanese phototypesetting. So one should not change the value of this dimension.

\jH (dimension) There is also a unit called "歯" which equals to 0.25 mm and used in Japanese phototypesetting. This \jH is the same \texttt{dimen} register as \jQ.

\ltj@dimen@zw (dimension) A temporal register for the “full-width” of current Japanese font. The command \zw sets this register to the correct value, and "return" this register itself.

\ltj@dimen@zh (dimension) A temporal register for the “full-height” (usually the sum of height of imaginary body and its depth) of current Japanese font. The command \zh sets this register to the correct value, and "return" this register itself.

\jfam (attribute) Current number of Japanese font family for math formulas.

\ltj@curjfnt (attribute) If this attribute is a positive number, it stores the font number of current Japanese font for horizontal direction. If this attribute is negative, it means that the Japanese font for horizontal direction is not loaded—Lua\TeX-ja only knows its size and JFM.

\ltj@curtfnt (attribute) Similar to \ltj@curjfnt, but with current Japanese font for vertical direction.

\ltj@charclass (attribute) The character class of a JAchar. This attribute is only set on a glyph_node which contains a JAchar.

\ltj@yablshift (attribute) The amount of shifting the baseline of alphabetic fonts in scaled point \((2^{-16}\text{ pt})\).

\ltj@ykb1shift (attribute) The amount of shifting the baseline of Japanese fonts in scaled point \((2^{-16}\text{ pt})\).

\ltj@tab1shift (attribute)
\ltj@tkb1shift (attribute)

\ltj@autospc (attribute) Whether the auto insertion of kanjiskip is allowed at the node.

\ltj@autoxspc (attribute) Whether the auto insertion of xkanjiskip is allowed at the node.

\ltj@icflag (attribute) An attribute for distinguishing "kinds" of a node. One of the following value is assigned to this attribute:

\textit{italic} (1) Kerns from italic correction (/\textbackslash{}), or from kerning information of a Japanese font. These kerns are “ignored” in the insertion process of J\texttt{Glue}, unlike explicit \textbackslash{}kern.

\textbf{packed} (2)

\textit{kinsoku} (3) Penalties inserted for the word-wrapping process (kinsoku shori) of Japanese characters.

\textit{from\_jfm}–(\textit{from\_jfm} + 63) (4–67) Glues/kerns from JFM.
\textit{kanji\_skip} (68), \textit{kanji\_skip\_jfm} (69) Glues from kanjiskip.
\textit{xkanji\_skip} (70), \textit{xkanji\_skip\_jfm} (71) Glues from xkanjiskip.
\textit{processed} (73) Nodes which is already processed by ....
ic\_processed (74) Glues from an italic correction, but already processed in the insertion process of JAglues.

boxbdd (75) Glues/kerns that inserted just the beginning or the ending of an hbox or a paragraph.

special\_jaglue (76) Glues from \(\texttt{\textbackslash insert[x]kanjiskip}\).

\textbackslash ltj@kat \(i\) (attribute) Where \(i\) is a natural number which is less than 7. These 7 attributes store bit vectors indicating which character block is regarded as a block of JAChars.

\textbackslash ltj@dir (attribute) dir\_node\_auto (128)

\textbackslash ltj@lineendcomment (counter)

Furthermore, LuaTEX-ja uses several user-defined whatsit nodes for internal processing. All those nodes except direction whatsits store a natural number (hence its type is 100). direction whatsits store a node list, hence its type is 110. Their user\_id (used for distinguish user-defined whatsits) are allocated by \texttt{luatexbase.newuserwhatsitid}.

inhibitglue Nodes for indicating that \texttt{\textbackslash inhibitglue} is specified. The value field of these nodes doesn’t matter.

stack\_marker Nodes for LuaTEX-ja’s stack system (see the next subsection). The value field of these nodes is current group level.

char\_by\_cid Nodes for JAchar which processes by luaotfload won’t be applied, and the character code is stored in the value field. Each node of this type are converted to a glyph\_node after processes by luaotfload. Nodes of this type is used in \texttt{\textbackslash CID} and \texttt{\textbackslash UTF}.

replace\_vs Similar to char\_by\_cid whatsits above. These nodes are for ALchar which the callback process of luaotfload won’t be applied.

begin\_par Nodes for indicating beginning of a paragraph. A paragraph which is started by \texttt{\textbackslash item} in list-like environments has a horizontal box for its label before the actual contents. So ...

direction

These whatsits will be removed during the process of inserting JAglues.

13.2 Stack system of LuaTEX-ja

Background LuaTEX-ja has its own stack system, and most parameters of LuaTEX-ja are stored in it. To clarify the reason, imagine the parameter kanjiskip is stored by a skip, and consider the following source:

\begin{verbatim}
1 \textbackslash ltjsetparameter{kanjiskip=0pt}ふがふが．%
2 \setbox0=\hbox{%
3 \textbackslash ltjsetparameter{kanjiskip=5pt}ほげほげ}
4 \box0．ぴよぴよ
\end{verbatim}

As described in Subsection 8.1, the only effective value of kanjiskip in an hbox is the latest value, so the value of kanjiskip which applied in the entire hbox should be 5pt. However, by the implementation method of LuaTEX, this “5 pt” cannot be known from any callbacks. In the tex/packaging.w, which is a file in the source of LuaTEX, there are the following codes:

\begin{verbatim}
void package(int c)
{
  scaled h; /* height of box */
  halfword p; /* first node in a box */
  scaled d; /* max depth */
  int grp;
  grp = cur_group;
}\end{verbatim}
d = box_max_depth;
unsave();
save_ptr -= 4;
if (cur_list.mode_field == ~hmode) {
  cur_box = filtered_hpack(cur_list.head_field,
      cur_list.tail_field, saved_value(1),
      saved_level(1), grp, saved_level(2));
  subtype(cur_box) = HLIST_SUBTYPE_HBOX;
}

Notice that unsave() is executed before filtered_hpack(), where hpack_filter callback is executed) here. So "5 pt" in the above source is orphaned at unsave(), and hence it can’t be accessed from hpack_filter callback.

■ Implementation The code of stack system is based on that in a post of Dev-luatex mailing list\[^1\].

These are two \TeX\ count registers for maintaining information: \ltj@@stack for the stack level, and \ltj@@group@level for the \TeX ‘s group level when the last assignment was done. Parameters are stored in one big table named charprop_stack_table, where charprop_stack_table[i] stores data of stack level i. If a new stack level is created by \ltjsetparameter, all data of the previous level is copied.

To resolve the problem mentioned in above paragraph “Background”, Lua\TeX-ja uses another trick. When the stack level is about to be increased, a whatsit node whose type, subtype and value are 44 (user defined), stack_marker and the current group level respectively is appended to the current list (we refer this node by stack_flag). This enables us to know whether assignment is done just inside a hbox.

Suppose that the stack level is \(s\) and the \TeX ‘s group level is \(t\) just after the hbox group, then:

- If there is no stack_flag node in the list of the contents of the hbox, then no assignment was occurred inside the hbox. Hence values of parameters at the end of the hbox are stored in the stack level \(s\).

- If there is a stack_flag node whose value is \(t + 1\), then an assignment was occurred just inside the hbox group. Hence values of parameters at the end of the hbox are stored in the stack level \(s + 1\).

- If there are stack_flag nodes but all of their values are more than \(t + 1\), then an assignment was occurred in the box, but it is done in more internal group. Hence values of parameters at the end of the hbox are stored in the stack level \(s\).

Note that to work this trick correctly, assignments to \ltj@@stack and \ltj@@group@level have to be local always, regardless the value of \globaldefs. To solve this problem, we use another trick: the assignment \directlua{tex.globaldefs=0} is always local.

13.3 Lua functions of the stack system

In this subsection, we will see how a user use Lua\TeX-ja’s stack system to store some data which obeys the grouping of \TeX.

The following function can be used to store data into a stack:
luatexja.stack.set_stack_table(index, <any> data)

Any values which except nil and NaN are usable as index. However, a user should use only negative integers or strings as index, since natural numbers are used by Lua\TeX-ja itself. Also, whether data is stored locally or globally is determined by luatexja.isglobal (stored globally if and only if luatexja.isglobal == ‘global’).

Stored data can be obtained as the return value of
luatexja.stack.get_stack_table(index, <any> default, <number> level)

where level is the stack level, which is usually the value of \ltj@@stack, and default is the default value which will be returned if no values are stored in the stack table whose level is level.

13.4 Extending Parameters

Keys for \ltjsetparameter and \ltjgetparameter can be extended, as in luatexja-adjust.

\protected\def\ltj@setpar@global{%  \relax\ifnum\globaldefs>0\directlua{luatexja.isglobal='global'}%  \else\directlua{luatexja.isglobal=''}\fi%
}\protected\def\ltjsetparameter#1{%  \ltj@setpar@global\setkeys[ltj]{japaram}{#1}\ignorespaces}
}\protected\def\ltjglobalsetparameter#1{%  \relax\ifnum\globaldefs<0\directlua{luatexja.isglobal=''}%  \else\directlua{luatexja.isglobal='global'}\fi%  \setkeys[ltj]{japaram}{#1}\ignorespaces}

Figure 12. Definition of parameter setting commands

Setting parameters Figure 12 shows the most outer definition of two commands, \texttt{\ltjsetparameter} and \texttt{\ltjglobalsetparameter}. Most important part is the last \texttt{\setkeys}, which is offered by the \texttt{xkeyval} package.

Hence, to add a key in \texttt{\ltjsetparameter}, one only have to add a key whose prefix is \texttt{ltj} and whose family is \texttt{japaram}, as the following.
\begin{verbatim}
\define@key[ltj]{japaram}{...}{...}
\end{verbatim}
\texttt{\ltjsetparameter} and \texttt{\ltjglobalsetparameter} automatically sets \texttt{luatexja.isglobal}. Its meaning is the following.
\begin{equation}
\text{luatexja.isglobal} = \begin{cases} 'global' & \text{(global assignment),} \\ '' & \text{(local assignment).} \end{cases}
\end{equation}

This is determined not only by command name (\texttt{\ltjsetparameter} or \texttt{\ltjglobalsetparameter}), but also by the value of \texttt{\globaldefs}.

Getting parameters \texttt{\ltjgetparameter} is implemented by a Lua script.

For parameters that do not need additional arguments, one only have to define a function in the table \texttt{luatexja.unary_pars}. For example, with the following function, \texttt{\ltjgetparameter{hoge}} returns a string \texttt{42}.
\begin{verbatim}
1 function luatexja.unary_pars.hoge (t)
2 return 42
3 end
\end{verbatim}
Here the argument of \texttt{luatexja.unary_pars.hoge} is the stack level of LuaTEX-ja's stack system (see Subsection 13.2).

On the other hand, for parameters that need an additional argument (this must be an integer), one have to define a function in \texttt{luatexja.binary_pars} first. For example,
\begin{verbatim}
1 function luatexja.binary_pars.fuga (c, t)
2 return tostring(c) .. ',', tostring(42)
3 end
\end{verbatim}
Here the first argument \texttt{t} is the stack level, as before. The second argument \texttt{c} is just the second argument of \texttt{\ltjgetparameter}.

For parameters that need an additional argument, one also have to execute the \TeX code like
\begin{verbatim}
\texttt{\ltj@@decl@array@param{fuga}}
\end{verbatim}
to indicate that "the parameter fuga needs an additional argument".

14 Linebreak after a Japanese Character
14.1 Reference: behavior in \TeX

In \TeX, a line break after a Japanese character doesn’t emit a space, since words are not separated by spaces in Japanese writings. However, this feature isn’t fully implemented in Lua\TeX-ja due to the specification of callbacks in Lua\TeX. To clarify the difference between \TeX and Lua\TeX, We briefly describe the handling of a line break in \TeX, in this subsection.

\TeX’s input processor can be described in terms of a finite state automaton, as that of \TeX in Section 2.5 of \cite{1}. The internal states are as follows:

• State $N$: new line
• State $S$: skipping spaces
• State $M$: middle of line
• State $K$: after a Japanese character

The first three states—$N$, $S$, and $M$—are as same as \TeX’s input processor. State $K$ is similar to state $M$, and is entered after Japanese characters. The diagram of state transitions are indicated in Figure 13. Note that \TeX doesn’t leave state $K$ after “beginning/ending of a group” characters.

14.2 Behavior in Lua\TeX-ja

States in the input processor of Lua\TeX is the same as that of \TeX, and they can’t be customized by any callbacks. Hence, we can only use process_input_buffer and token_filter callbacks for to suppress a space by a line break which is after Japanese characters.

However, token_filter callback cannot be used either, since a character in category code 5 (end-of-line) is converted into an space token in the input processor. So we can use only the process_input_buffer callback. This means that suppressing a space must be done just before an input line is read.

Considering these situations, handling of an end-of-line in Lua\TeX-ja are as follows:

A character whose character code is $\texttt{\backslash t j lineendcomment}$ is appended to an input line, before Lua\TeX actually process it, if and only if the following three conditions are satisfied:

\footnote{Its default value is “FFFFF, so $\texttt{U+FFFFF}$ is used. The category code of $\texttt{U+FFFFF}$ is set to 14 (comment) by Lua\TeX-ja.}
1. The category code of \endlinechar is 5 (end-of-line).
2. The category code of \ltjlineendcomment itself is 14 (comment).
3. The input line matches the following "regular expression":
   \begin{verbatim}
   (any char)\(\text{JAchar}\)\{\text{catcode} = 1\} \cup \{\text{catcode} = 2\}\end{verbatim}

\begin{remark}
The following example shows the major difference from the behavior of p\TeX.
\begin{verbatim}
\fontspec[ Ligatures=TeX]{Linux Libertine O}
\setparameter{autoxspacing=false}
\setparameter{jacharrange={-6}}xあ
y\setparameter{jacharrange={+6}}zい
u
\end{verbatim}
\end{remark}

It is not strange that “あ” does not printed in the above output. This is because \TeX Gyre Termes does not contain “あ”, and because “あ” in line 3 is considered as an Alchar.

Note that there is no space before “y” in the output, but there is a space before “u”. This follows from following reasons:

- When line 3 is processed by process_input_buffer callback, “あ” is considered as an JAchar. Since line 3 ends with an JAchar, the comment character (whose character code is \ltjlineendcomment) is appended to this line, and hence the linebreak immediately after this line is ignored.
- When line 4 is processed by process_input_buffer callback, “い” is considered as an Alchar. Since line 4 ends with an Alchar, the linebreak immediately after this line emits a space.

## 15 Patch for the listings Package

It is well-known that the listings package outputs weird results for Japanese input. The listings package makes most of letters active and assigns output command for each letter ([2]). But Japanese characters are not included in these activated letters. For p\TeX series, there is no method to make Japanese characters active; a patch \jstyling.sty ([4]) resolves the problem forcibly.

In Lua\TeX-ja, the problem is resolved by using the process_input_buffer callback. The callback function inserts the output command (active character \ltjlineendcomment) before each letter above U+0080. This method can omits the process to make all Japanese characters active (most of the activated characters are not used in many cases).

If the listings package and Lua\TeX-ja were loaded, then the patch \ltjp-listings is loaded automatically at \begin{document}.

### 15.1 Notes and additional keys

\begin{variationselectors}
\ltjp-listings add two keys, namely \texttt{vsraw} and \texttt{vscmd}, which specify how variation selectors are treated in \texttt{lstlisting} or other environments. Note that these additional keys are not usable in the preamble, since \ltjp-listings is loaded at \begin{document}.

\texttt{vsraw} is a key which takes a boolean value, and its default value is false.

- If the \texttt{vsraw} key is true, then variation selectors are "combined" with the previous character.

\begin{verbatim}
\begin{lstlisting}[vsraw=true]
すべて,ほとんどの,非常に
\end{lstlisting}
\end{verbatim}

- If the \texttt{vsraw} key is false, then variation selectors are typeset by an appropriate command, which is specified by the \texttt{vscmd} key. The default setting of the \texttt{vscmd} key produces the following.

\begin{verbatim}
\begin{lstlisting}
すべて,ほとんど,非常に
\end{lstlisting}
\end{verbatim}

\end{variationselectors}

\end{document}
For example, the following code is the setting of the vscmd key in this document.

```latex
\def\IVSA#1#2#3#4#5{\hbox to 1em{\hss\textcolor{blue}{\raisebox{3.5pt}{\normalfont\ttfamily\fboxsep=0.5pt\fbox{\hbox to 0.75em{\hss\tiny \oalign{0#1#2\crcr#3#4#5\crcr}\hss}}}}\hss}}
\lstset{vscmd=\IVSB}
```

The default output command of variation selectors is stored in \ltjlistingsvsstdcmd.

The doubleletterspace key  

Even the column format is \texttt{c} fixed, sometimes characters are not vertically aligned. The following example is typeset with basewidth=2em, and you’ll see the leftmost ‘H’ are not vertically aligned.

```
\begin{verbatim}
\texttt{H :}
\texttt{H H H H :}
\end{verbatim}
```

lltj-listing adds the doubleletterspace key (not activated by default, for compatibility) to improve the situation, namely doubles inter-character space in each output unit. With this key, the above input now produces better output.

```
\begin{verbatim}
\texttt{H :}
\texttt{H H H H :}
\end{verbatim}
```

15.2 Class of characters

Roughly speaking, the listings package processes input as follows:

1. Collects \textit{letters} and \textit{digits}, which can be used for the name of identifiers.
2. When reading an \textit{other}, outputs the collected character string (with modification, if needed).
3. Collects \textit{others}.
4. When reading a \textit{letter} or a \textit{digit}, outputs the collected character string.
5. Turns back to 1.

By the above process, line breaks inside of an identifier are blocked. A flag \texttt{\lst@ifletter} indicates whether the previous character can be used for the name of identifiers or not.

For Japanese characters, line breaks are permitted on both sides except for brackets, dashes, etc. Hence the patch lltj-listings introduces a new flag \texttt{\lst@ifkanji}, which indicates whether the previous character is a Japanese character or not. For illustration, we introduce following classes of characters:

<table>
<thead>
<tr>
<th>Letter</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{\lst@ifletter}</td>
<td>T</td>
</tr>
<tr>
<td>\texttt{\lst@ifkanji}</td>
<td>F</td>
</tr>
</tbody>
</table>

| Meaning | char in an identifier | other alphabet |

<table>
<thead>
<tr>
<th>Kanji</th>
<th>Open</th>
<th>Close</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{\lst@ifletter}</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>\texttt{\lst@ifkanji}</td>
<td>T</td>
<td>F</td>
</tr>
</tbody>
</table>

| Meaning | most of Japanese char | opening brackets | closing brackets |
Note that *digits* in the listings package can be Letter or Other according to circumstances.

For example, let us consider the case an Open comes after a Letter. Since an Open represents Japanese open brackets, it is preferred to be permitted to insert line break after the Letter. Therefore, the collected character string is output in this case.

The following table summarizes $5 \times 5 = 25$ cases:

<table>
<thead>
<tr>
<th>Prev</th>
<th>Next</th>
<th>Letter</th>
<th>Other</th>
<th>Kanji</th>
<th>Open</th>
<th>Close</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanji</td>
<td>Letter</td>
<td>collects</td>
<td>outputs</td>
<td>collects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td>Other</td>
<td>outputs</td>
<td>collects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close</td>
<td></td>
<td></td>
<td>collects</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the above table,

- “outputs” means to output the collected character string (i.e., line breaking is permitted there).
- “collects” means to append the next character to the collected character string (i.e., line breaking is prohibited there).

Characters above or equal to $\text{U+0080}$ except *Variation Selectors* are classified into above 5 classes by the following rules:

- **ALchar**s above or equal to $\text{U+0080}$ are classified as Letter.
- **JAchar**s are classified in the order as follows:
  1. Characters whose `prebreakpenalty` is greater than or equal to 0 are classified as Open.
  2. Characters whose `postbreakpenalty` is greater than or equal to 0 are classified as Close.
  3. Characters that don’t satisfy the above two conditions are classified as Kanji.

The width of halfwidth kana ($\text{U+FF61}$–$\text{U+FF9F}$) is same as the width of **ALchar**; the width of the other **JAchar**s is double the width of **ALchar**.

This classification process is executed every time a character appears in the `\lstlisting` environment or other environments/commands.

### 16 Cache Management of LuaTEx-ja

**LuaTEx-ja** creates some cache files to reduce the loading time. in a similar way to the `luaotfload` package:

- Cache files are usually stored in (and loaded from) `$\text{TEXMFVAR/luatexja/}$`
- In addition to caches of the text form (the extension is “.lua.gz”, because they are compressed by `gzip`), caches of the *binary* (bytecode) form are supported.
  - In loading a cache, the binary cache precedes the text form.
  - When LuaTEx-ja updates a compressed text cache `hoge.lua.gz`, its binary version is also updated.

#### 16.1 Use of cache

**LuaTEx-ja** uses the following cache:

`\ttj-cid-auto-adobe-japan1.{lua.gz,1uc}`

The font table of a CID-keyed non-embedded Japanese font. This is loaded in every run. It is created from three CMaps, `UniJIS2004-UTF32-{H,V}` and `Adobe-Japan1-UCS2`, and this is why these two CMaps are needed in the first run of LuaTEx-ja.

Similar caches are created as Table 16, if you specified `cid` key in `\jfont` to use other CID-keyed non-embedded fonts for Chinese or Korean, as in Page 26.
Table 16. cid key and corresponding files

<table>
<thead>
<tr>
<th>cid key</th>
<th>name of the cache</th>
<th>used CMaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe-Japan1-*</td>
<td>1tj-cid-auto-adobe-japan1.{lua.gz,luc}</td>
<td>UniJIS2004-UTF32-*</td>
</tr>
<tr>
<td>Adobe-Korea1-*</td>
<td>1tj-cid-auto-adobe-korea1.{lua.gz,luc}</td>
<td>UniKS-UTF32-*</td>
</tr>
<tr>
<td>Adobe-KR-*</td>
<td>1tj-cid-auto-adobe-kr.{lua.gz,luc}</td>
<td>UniAKR-UTF32-*</td>
</tr>
<tr>
<td>Adobe-GB1-*</td>
<td>1tj-cid-auto-adobe-gb1.{lua.gz,luc}</td>
<td>UniGB-UTF32-*</td>
</tr>
<tr>
<td>Adobe-CNS1-*</td>
<td>1tj-cid-auto-adobe-cns1.{lua.gz,luc}</td>
<td>UniCNS-UTF32-*</td>
</tr>
</tbody>
</table>

ltj-kinsoku.luc
The bytecode cache which default kinsoku parameters are stored.

ltj-jisx0208.luc
The bytecode version of ltj-jisx0208.lua. This is the conversion table between JIS X 0208 and Unicode which is used in Kanji-code conversion commands for compatibility with \TeX{}.

ltj-ivd_aj1.luc
The bytecode version of ltj-ivd_aj1.lua.

extra_***.{lua.gz,luc}
This file contains some information (especially for vertical typesetting) about the font '***'.

16.2 Internal

Cache management system of LuaTeX-ja is stored in luatexja.base (ltj-base.lua). There are four public functions for cache management in luatexja.base, where ⟨filename⟩ stands for the file name without suffix:

save_cache(⟨filename⟩, ⟨data⟩)
Save a non-nil table ⟨data⟩ into a cache ⟨filename⟩. Both the compressed text form ⟨filename⟩.lua.gz and its binary version are created or updated.

save_cache_luc(⟨filename⟩, ⟨data⟩, ⟨serialized_data⟩)
Same as save_cache, except that only the binary cache is updated. The third argument ⟨serialized_data⟩ is not usually given. But if this is given, it is treated as a string representation of ⟨data⟩.

load_cache(⟨filename⟩, ⟨outdate⟩)
Load the cache ⟨filename⟩. ⟨outdate⟩ is a function which takes one argument (the contents of the cache), and its return value is whether the cache is outdated.

load_cache first tries to read the binary cache ⟨filename⟩.luc. If its contents is up-to-date, load_cache returns the contents. If the binary cache is not found or its contents is outdated, load_cache tries to read the compressed text form ⟨filename⟩.lua.gz. Hence, the return value of load_cache is non-nil, if and only if the updated cache is found.

remove_cache(⟨filename⟩)
Remove the cache ⟨filename⟩.
References


