The LuaTeX-ja package

The LuaTeX-ja project team

20200209.0 (February 9, 2020)
Contents

I User’s manual 4

1 Introduction 4
  1.1 Backgrounds .................................................. 4
  1.2 Major changes from \TeX ................................. 4
  1.3 Notations ..................................................... 5
  1.4 About the project ......................................... 6

2 Getting Started 7
  2.1 Installation .................................................. 7
  2.2 Cautions ...................................................... 8
  2.3 Using in plain \TeX ...................................... 8
  2.4 Using in \LaTeX ........................................... 9

3 Changing Fonts 10
  3.1 plain \TeX and \LaTeX .................................. 10
  3.2 lualatexja-fontspec package ............................ 11
  3.3 Presets of Japanese fonts .............................. 12
  3.4 \CID, \UTF , and macros in japanese-otf package ......... 12

4 Changing Internal Parameters 12
  4.1 Range of JAchars ........................................ 12
  4.2 \kanjiskip and xkanjiskip ................................ 15
  4.3 Insertion setting of xkanjiskip ......................... 15
  4.4 Shifting the baseline .................................... 16
  4.5 kinsoku parameters and OpenType features .......... 17

II Reference 18

5 catcode in Lua\TeX-ja 18
  5.1 Preliminaries: \textbackslash catcode in \TeX and up\TeX ..... 18
  5.2 Case of Lua\TeX-ja ...................................... 18
  5.3 Non-kanji characters in a control word ............. 18

6 Directions 19
  6.1 Boxes in different direction ............................ 19
  6.2 Getting current direction ............................... 21
  6.3 Overridden box primitives .............................. 22

7 Font Metric and Japanese Font 22
  7.1 \texttt{jfont} ........................................... 22
  7.2 \texttt{tfont} ........................................... 24
  7.3 Default Japanese fonts and JFMs ....................... 25
  7.4 Prefix psft ............................................. 25
  7.5 Structure of a JFM file ................................ 26
  7.6 Math font family ....................................... 30
  7.7 Callbacks .................................................. 30
8 Parameters
8.1 \ltjsetparameter .................................................. 32
8.2 \ltjgetparameter .................................................. 34
8.3 Alternative Commands to \ltjsetparameter ...................... 35

9 Other Commands for plain \TeX and \LaTeXe 
9.1 Commands for compatibility with p\TeX .................................. 35
9.2 inhibitglue .......................................................... 36
9.3 \ltjfakeboxbdd, \ltjfakeparbegin .................................. 36
9.4 \ltjdeclarealtfont .................................................. 36

10 Commands for \LaTeXe 
10.1 Loading Japanese fonts in \LaTeXe ..................................... 37
10.2 Patch for NFSS2 .................................................... 37
10.3 Detail of \fontfamily command ..................................... 39
10.4 Notes on \DeclareTextSymbol ..................................... 40
10.5 \strutbox ................................................................ 40

11 Addon packages
11.1 luatexja-fontspec ................................................... 40
11.2 luatexja-otf ......................................................... 42
11.3 luatexja-adjust ...................................................... 43
11.4 luatexja-ruby ........................................................ 43
11.5 \ltxjext.sty .......................................................... 44
11.6 luatexja-preset ...................................................... 44
11.6.1 General Options .................................................. 45
11.6.2 Presets which support multi weights ......................... 46
11.6.3 Presets which do not support multi weights .................. 48
11.6.4 Presets which use HG fonts ................................... 49
11.6.5 Define/Use Custom Presets ................................... 49

III Implementations

12 Storing Parameters
12.1 Used dimensions, attributes and whatsit nodes ............... 51
12.2 Stack system of Lua\TeX-ja ......................................... 52
12.3 Lua functions of the stack system ................................ 53
12.4 Extending Parameters .............................................. 53

13 Linebreak after a Japanese Character
13.1 Reference: behavior in p\TeX ..................................... 54
13.2 Behavior in Lua\TeX-ja .............................................. 55

14 Patch for the listings Package
14.1 Notes and additional keys ......................................... 56
14.2 Class of characters ................................................ 57

15 Cache Management of Lua\TeX-ja
15.1 Use of cache ......................................................... 58
15.2 Internal ............................................................... 59
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 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`ぁ=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='ぁ',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.¹

■ 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 13.¹

¹LuaTEX-ja 20200127.0 introduces \ltj@zw and \ltj@zh, which are copy of \zw and \zh.
**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 Lua\TeX’s internal ligature handling is *node-based* (e.g., \texttt{of}{}fice 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., \texttt{\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 (\texttt{hbox}) between it instead:
  ちょっと/ひとつ

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

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

\textbf{\textbackslash discretionary} Japanese characters in discretionary break (\texttt{\discretionary}) is not supported.

**Greek and Cyrillic letters, and ISO 8859-1 symbols** By default, Lua\TeX-ja uses Japanese fonts to typeset Greek and Cyrillic letters. To change this behavior, put \texttt{\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. If you are using \tex prior to 2017/01/01, these characters are not typeset correctly without the \texttt{\fontspec} (and \texttt{luatexja-fontspec}) package.

### 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 Hiragana, 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 \texttt{prebreakpenalty}) means an internal parameter for Japanese typesetting, and it is used as a key in \texttt{\ltjsetparameter} command.

- A word in a sens-serif font without underline (like \texttt{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 \texttt{Lua\TeX\-ja} package.

- \texttt{Lua\TeX} 1.07.0 (or later)
- recent \texttt{luaotfload} (v3.1 or later recommended)
- adobemapping (Adobe cmap and pdfmapping files)
- \texttt{etoolbox}, \texttt{everyseq} (if you want to use \texttt{Lua\TeX\-ja} with \LaTeX\ \texttt{2\epsilon})
- \texttt{ltxcmds}, \texttt{pdftexcmds}, \texttt{filehook}, \texttt{atbegshi}
- \texttt{fontspec} v2.7c (or later)
- IPAex fonts (\url{http://ipafont.ipa.go.jp/})

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

- MiK\TeX\ (in \texttt{luatexja.tar.lzma}); see the next subsection
- \TeX\ Live (in \texttt{texmf-dist/tex/luatex/luatexja})
- W32\TeX\ (in \texttt{luatexja.tar.xz})

IPAex fonts are also available in these distributions.

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

Especially, \emph{We don't recommend defining a Japanese font with HarfBuzz, by specifying \texttt{Renderer=Harfbuzz} etc. (\texttt{fontspec}) or \texttt{mode=harf} (otherwise)}.

\textbf{Manual installation}

1. Download the source, by one of the following method. At the present, \texttt{Lua\TeX\-ja} has no \textit{stable} release.

   - Clone the Git repository by
     
     $\text{git clone git://git.osdn.jp/gitroot/latex\-ja/luatexja.git}$

   - Download the \texttt{tar.gz} archive of HEAD in the master branch from
     
     \url{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 \texttt{src/} and several other sub-directories. But only the contents in \texttt{src/} are needed to work \texttt{Lua\TeX\-ja}.

3. If you downloaded this package from CTAN, you have to run following commands to generate classes and \texttt{ltj\-kinsoku.lua} (the file which stores default \textit{kinsoku} parameters):

   $\text{cd src}$
   $\text{lualatex ltjclasses.ins}$
   $\text{lualatex ltjsclasses.ins}$
   $\text{lualatex ltjltxdoc.ins}$
   $\text{luatex ltj\-kinsoku\_make.tex}$
Do not forget processing `ltj-kinsoku_make.tex` and `ltj-kinsoku_make.tex` used here are not needed in regular use.

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 `mktexlsr` is needed to update the file name database, make it so.

2.2 Cautions

For changes from `pTeX`, 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.
- `LuaTeX-ja` is very slower than `pTeX`, and uses a lot of memory.
- **(Outdated) note for MiKTeX users**  `LuaTeX-ja` requires that several CMap files\(^2\) must be found from `LuaTeX`. Strictly speaking, those CMaps are needed only in the first run of `LuaTeX-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 `LuaTeX-ja` in this directory, and finally delete the temporary directory.

- Note that when `LuaTeX-ja` is loaded in plain `TeX` nor `MiKTeX 2019-10-01`, we cannot use color specification on font loading, such as

```
\font\hoge=lmroman10-regular.otf:color=FF0000 % \font primitive
\DeclareFontShape{TU}{...}{...}{...}{<-> lmsans10-boldoblique:color=005599}{} % NFSS
\setmainfont{texgyretermes}[Color=220022] % fontspec
```

This is because codes for shifting baseline in math mode (`LuaTeX-ja`) collide with and prevents loading codes for font color (luaotfload) in these environments. *We recommend to use `MiKTeX 2020-02-02`, since we can avoid this collision in `MiKTeX 2020-02-02`.*

2.3 Using in plain TeX

To use `LuaTeX-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><strong>yoko</strong> (horizontal)</td>
<td><strong>mincho</strong> gothic</td>
<td>IPAex Mincho IPAex Gothic</td>
<td>\textmin</td>
<td>\sevenmin</td>
<td>\fivemin</td>
</tr>
<tr>
<td><strong>tate</strong> (vertical)</td>
<td><strong>mincho</strong> gothic</td>
<td>IPAex Mincho IPAex Gothic</td>
<td>\text tengt</td>
<td>\seventengt</td>
<td>\fivegt</td>
</tr>
</tbody>
</table>

\(^2\)UniJIS2004-UTF32-\{(H,V) and Adobe--Japan1--UCS2.\)
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
  
  \[ (0.25 \cdot 0.962216 \cdot 10 \text{pt})^{+1\text{pt}} - 1\text{pt} = 2.40554 \text{pt}^{+1\text{pt}}. \]

2.4 Using in \LaTeX

Using in \LaTeX\ was basically same. To set up the minimal environment for Japanese, you only have to load \texttt{luatexja.sty}:

\begin{verbatim}
\usepackage{luatexja}
\end{verbatim}

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

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

- Traditionally, Japanese documents use only two families: \textit{mincho} (明朝体) and \textit{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 \texttt{\jttdefault} specifies the Japanese font family in \texttt{\verb} or \texttt{\verbatim} environment, and its default value is \texttt{\mcfamily} (\textit{mincho} family).\textsuperscript{3} \texttt{\LaTeX\-ja} does not define commands to only switch current Japanese font family to \texttt{\jttdefault}.

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

\begin{verbatim}
\begin{array}{ccc}
\hline
\textbf{classification} & \textbf{commands} & \textbf{family} \\
\hline
mincho (明朝体) & \texttt{\textmc{...} \{\mcfamily ...\}} & \texttt{\mcdefault} \\
gothic (ゴシック体) & \texttt{\textgt{...} \{\gtfamily ...\}} & \texttt{\gtdefault} \\
(Japanese counterpart for typewriter font) & - & \texttt{\jttdefault} \\
\hline
\end{array}
\end{verbatim}

Note that the bold series (series \texttt{bx} or \texttt{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 \texttt{disablejfam} option at loading \texttt{\LaTeX\-ja}. This option prevents loading a patch for \LaTeX, which are needed to support Japanese characters in math mode. Without \texttt{disablejfam} option, one can typeset Japanese characters in math mode as \texttt{\$\text{\text{	extAA}}\$} (see Page 10) as before. Japanese characters in math mode are typeset by the font family \texttt{mc}.

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

\begin{verbatim}
\renewcommand{\kanjifamilydefault}{\gtdefault}
\end{verbatim}

However, above settings are not sufficient for Japanese-based documents. To typeset Japanese-based documents, you are better to use class files other than \texttt{article.cls}, \texttt{book.cls}, and so on. At the present, \texttt{\LaTeX\-ja} has the counterparts of \texttt{jclasses} (standard classes in \TeX\) and \texttt{jsclasses} (classes by Haruhiko Okumura), namely, \texttt{ltjclasses}\textsuperscript{4} and \texttt{ltjsclasses}\textsuperscript{5}.

\textsuperscript{3}When \texttt{ltjsclasses} classes are used, or \texttt{luatexja-fontspec} (or \texttt{luatexja-presets}) is loaded with \texttt{match} option, \texttt{\jttfamily} changes the current Japanese font amily to \texttt{\jttdefault}. These classes and packages also redefine \texttt{\jttdefault} to \texttt{\gtdefault} (gothic family).

\textsuperscript{4}\texttt{ltjarticle.cls}, \texttt{ltjbook.cls}, \texttt{ltjreport.cls}, \texttt{ltjsarticle.cls}, \texttt{ltjsbook.cls}, \texttt{ltjsreport.cls}, \texttt{ltjskiyou.cls}, \texttt{ltjreport.cls}, \texttt{ltjtarticle.cls}, \texttt{ltjtbook.cls}, \texttt{ltjtreport.cls}. The latter \texttt{ltjt*.cls} are for vertically written Japanese documents.

\textsuperscript{5}\texttt{ltjsarticle.cls}, \texttt{ltjsbook.cls}, \texttt{ltjsreport.cls}, \texttt{ltjskiyou.cls}. 


Original \texttt{jsclasses} use \texttt{\mag} primitive to set the main document font size. However, \LaTeX{} does not support \texttt{\mag} in PDF output, so \texttt{ltjsclasses} uses the \texttt{nomag*} option by default to set the main font size. If this causes some unexpected behavior, specify \texttt{nomag} option in \texttt{\documentclass}.

\begin{itemize}
\item \textbf{geometry package and classes for vertical writing} It is well-known that the \texttt{geometry} package produces the following error, when classes for vertical writing is used:

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

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

3 Changing Fonts

3.1 \tex and \LaTeX{} $\varepsilon$

\begin{itemize}
\item \textbf{plain \texttt{\LaTeX}} To change Japanese fonts in plain \texttt{\LaTeX}, you must use the command \texttt{\jfont} and \texttt{\tfont}. So please see Subsection 7.1.
\item \textbf{\LaTeX{} $\varepsilon$ (NFSS2)} For \texttt{\LaTeX{} $\varepsilon$}, \LaTeXX-ja adopted most of the font selection system of \texttt{p\LaTeX{} $\varepsilon$} (in \texttt{plfonts.dtx}).
\end{itemize}

\begin{itemize}
\item Commands \texttt{\fontfamily}, \texttt{\fontseries}, and \texttt{\fontshape} can be used to change attributes of Japanese fonts.
\end{itemize}

\begin{table}[h]
\centering
\begin{tabular}{l|llll}
\hline
\textbf{encoding} & \textbf{family} & \textbf{series} & \textbf{shape} & \textbf{selection} \\
\hline
Alphabetic fonts & \texttt{\romanencoding} & \texttt{\romanfamily} & \texttt{\romanseries} & \texttt{\romanshape} & \texttt{\useroman} \\
Japanese fonts & \texttt{\kanjiencoding} & \texttt{\kanjifamily} & \texttt{\kanjiseries} & \texttt{\kanjishape} & \texttt{\usekanji} \\
both & - & - & \texttt{\fontseries} & \texttt{\fontshape} & - \\
auto select & \texttt{\fontencoding} & \texttt{\fontfamily} & - & \texttt{\fontshape} & - \\
\hline
\end{tabular}
\caption{Font Selection System in \texttt{\LaTeX{} $\varepsilon$}}
\end{table}

\begin{verbatim}
\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 both. For the detail, see Subsection 10.2.
\end{verbatim}

\begin{itemize}
\item For defining a Japanese font family, use \texttt{\DeclareKanjiFamily} instead of \texttt{\DeclareFontFamily}. (In previous version of \LaTeXX-ja, using \texttt{\DeclareFontFamily} didn't cause any problem. But this no longer applies the current version.)
\item Defining a Japanese font shape can be done by usual \texttt{\DeclareFontShape}:
\begin{verbatim}
\DeclareFontShape{JY3}{mc}{b}{n}{<-> s*KozMinPr6N-Bold:jfm=ujis;-kern}{% Kozuka Mincho Pr6N Bold
\end{verbatim}
\end{itemize}

\begin{itemize}
\item \textbf{Japanese characters in math mode} Since \texttt{p\LaTeX} supports Japanese characters in math mode, there are sources like the following:
\begin{verbatim}
1. $f_\text{高温}$ (\textit{f} \text{high temperature}).
2. \[ y=(x-1)^2+2 \text{ は} y>0 ]
3. $5 \in \text{素} : = \{ p \in \mathbb{N} : \text{p is a prime} \}.$
\end{verbatim}
\end{itemize}

\begin{verbatim}
\$\text{高温}\$ (\textit{f} \text{high temperature}).
\end{verbatim}

\begin{verbatim}
\$y=(x-1)^2+2 \text{ は} y>0$
\end{verbatim}

\begin{verbatim}
5 \in \text{素} : = \{ p \in \mathbb{N} : \text{p is a prime} \}.
\end{verbatim}

\footnote{Same effect as the BXjacls classes (by Takayuki Yato) and jsclasses. However, these classes uses only \texttt{\LaTeX} code, but \texttt{ltjsclasses} uses Lua code.}
Table 1. Commands of luatexja-fontspec

<table>
<thead>
<tr>
<th>Japanese fonts</th>
<th>\jfontspec</th>
<th>\setmainjfont</th>
<th>\setsansjfont</th>
<th>\setmonojfont</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphabetic fonts</td>
<td>\fontspec</td>
<td>\setmainfont</td>
<td>\setsansfont</td>
<td>\setmonofont</td>
</tr>
<tr>
<td>Japanese fonts</td>
<td>\newjfontfamily</td>
<td>\renewjfontfamily</td>
<td>\setjfontfamily</td>
<td></td>
</tr>
<tr>
<td>Alphabetic fonts</td>
<td>\newfontfamily</td>
<td>\renewfontfamily</td>
<td>\setfontfamily</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>

We (the project members of LuaTeX-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 “よって” 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. ($f_{\text{high temperature}}$).
3. $y = (x-1)^2 + 2 \quad \text{よって} \quad y > 0$
4. $5 \in \text{素} := \{ p \in \mathbb N : p \text{ is a prime}\}$.

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 LuaTeX-ja is loaded with \DisableJFont option, one cannot write Japanese characters in math mode as $素$. At that case, one have to use \mbox (or \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:

\texttt{match}

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

\texttt{pass=⟨opts⟩}

\texttt{(Obsoleted)} Specify options ⟨opts⟩ which will be passed to the fontspec package.

\texttt{scale=⟨float⟩}

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 \texttt{\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 \texttt{\Cjascale} is not defined.

\texttt{\Cjascale} is defined in \texttt{ltjclasses} and \texttt{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}
\usepackage[no-math]{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.

\begin{verbatim}
\jfontspec[CJKShape=NLC]{KozMinPr6N-Regular} \jfontspec[CJKShape=JIS1990]{KozMinPr6N-Regular}
\end{verbatim}

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 Subsecion 11.6. The following presets are defined:

- haranoaji
- hiragino-pro
- hiragino-pron
- ipa
- ipa-hg
- ipaex
- ipaex-hg
- kozuka-pr6
- kozuka-pr6n
- kozuka-pro
- moga-mobo
- moga-mobo-ex
- bizud
- morisawa-pr6n
- 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[kozuka-pr6n]{luatexja-preset}
which means that Kozuka Pr6N 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.

\begin{verbatim}
\jfontspec{KozMinPr6N-Regular.otf}
森\UTF{9DD7}外と内田百\UTF{9592}とが\UTF{9AD9}
島屋を行 く。
\end{verbatim}

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.

4.1 Range of JAchars

Lua\TeX-ja divides the Unicode codespace U+0080–U+1FFFFF into character ranges, numbered 1 to 217. The grouping can be (globally) customized by \ltjdefcharrange. The next line adds whole characters in Supplementary Ideographic Plane and the character “漢” to the character range 100.
\ltjdefcharrange{100}{“20000-”2FFFF,”漢}
Table 2. Characters in predefined character range 8.

<table>
<thead>
<tr>
<th>Character</th>
<th>Unicode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>§</td>
<td>U+00A7</td>
<td>Section Sign</td>
</tr>
<tr>
<td>'</td>
<td>U+00B0</td>
<td>Degree sign</td>
</tr>
<tr>
<td>´</td>
<td>U+00B4</td>
<td>Spacing acute</td>
</tr>
<tr>
<td>×</td>
<td>U+00D7</td>
<td>Multiplication sign</td>
</tr>
<tr>
<td>°</td>
<td>U+00B0</td>
<td>Degree sign</td>
</tr>
<tr>
<td>±</td>
<td>U+00B1</td>
<td>Plus-minus sign</td>
</tr>
<tr>
<td>¶</td>
<td>U+00B6</td>
<td>Paragraph sign</td>
</tr>
<tr>
<td>÷</td>
<td>U+00F7</td>
<td>Division Sign</td>
</tr>
</tbody>
</table>

Table 3. Unicode blocks in predefined character range 1.

<table>
<thead>
<tr>
<th>Block Name</th>
<th>Start</th>
<th>End</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latin-1 Supplement</td>
<td>U+0080</td>
<td>U+00FF</td>
<td></td>
</tr>
<tr>
<td>Latin-Extended-B</td>
<td>U+0100</td>
<td>U+017F</td>
<td></td>
</tr>
<tr>
<td>IPA Extensions</td>
<td>U+0180</td>
<td>U+024F</td>
<td></td>
</tr>
<tr>
<td>Spacing Modifier Letters</td>
<td>U+0250</td>
<td>U+02AF</td>
<td></td>
</tr>
<tr>
<td>Latin Extended Additional</td>
<td>U+0300</td>
<td>U+036F</td>
<td></td>
</tr>
<tr>
<td>Latin-Extended-A</td>
<td>U+0100</td>
<td>U+017F</td>
<td></td>
</tr>
<tr>
<td>IPA Extensions</td>
<td>U+0180</td>
<td>U+024F</td>
<td></td>
</tr>
<tr>
<td>Spacing Modifier Letters</td>
<td>U+0250</td>
<td>U+02AF</td>
<td></td>
</tr>
</tbody>
</table>

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 ALchar and 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 ALchar,
- a character which belongs character ranges 2, 3, 6, 7, and 9 is JAchar.

\ltjsetparameter{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 ALchar", and positive integer \(+n\) means that "... is a JAchar".

Note that characters U+0000–U+007F are always treated as an 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 JAchar or not by default. These settings are similar to the 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\(^1\) 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\(^1\) 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\(^1\) Greek and Cyrillic letters. JIS X 0208 (hence most of Japanese fonts) has some of these characters.

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

Range 3\(^1\) Miscellaneous symbols. The block list is indicated in Table 4.

Range 9\(^1\) The intersection of the “General Punctuation” block (U+2000–U+206F) and Adobe-Japan1-7 character collection. This character range characters in Table 5.
Table 4. Unicode blocks in predefined character range 3.

<table>
<thead>
<tr>
<th>Range</th>
<th>Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>U+2800-U+28FF Superscripts and Subscripts</td>
</tr>
<tr>
<td></td>
<td>U+28A0-U+28FF Currency Symbols</td>
</tr>
<tr>
<td></td>
<td>U+2100-U+214F Letterlike Symbols</td>
</tr>
<tr>
<td></td>
<td>U+2190-U+21FF Arrows</td>
</tr>
<tr>
<td></td>
<td>U+2300-U+23FF Miscellaneous Technical</td>
</tr>
<tr>
<td></td>
<td>U+2500-U+257F Box Drawing</td>
</tr>
<tr>
<td></td>
<td>U+255A-U+25FF Geometric Shapes</td>
</tr>
<tr>
<td></td>
<td>U+2700-U+27FF Dingbats</td>
</tr>
<tr>
<td></td>
<td>U+2900-U+29FF Misc. Math Symbols-B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Range</th>
<th>Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>U+2800-U+28FF Comb. Diacritical Marks for Symbols</td>
</tr>
<tr>
<td></td>
<td>U+2800-U+28FF Number Forms</td>
</tr>
<tr>
<td></td>
<td>U+2200-U+22FF Mathematical Operators</td>
</tr>
<tr>
<td></td>
<td>U+2150-U+215F Control Pictures</td>
</tr>
<tr>
<td></td>
<td>U+25A0-U+25FF Block Elements</td>
</tr>
<tr>
<td></td>
<td>U+2600-U+26FF Miscellaneous Symbols</td>
</tr>
<tr>
<td></td>
<td>U+2900-U+29FF Supplemental Arrows-B</td>
</tr>
<tr>
<td></td>
<td>U+2800-U+28FF Misc. Symbols and Arrows</td>
</tr>
</tbody>
</table>

Table 5. Characters in predefined character range 9.

<table>
<thead>
<tr>
<th>Range</th>
<th>Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Hot double low-9 quotation mark</td>
</tr>
<tr>
<td></td>
<td>Double dagger</td>
</tr>
<tr>
<td></td>
<td>Single left-pointing angle quot.</td>
</tr>
<tr>
<td></td>
<td>Overline</td>
</tr>
<tr>
<td></td>
<td>Exclamation question mark</td>
</tr>
<tr>
<td></td>
<td>Variation Selectors Supp.</td>
</tr>
</tbody>
</table>

Table 6. Unicode blocks in predefined character range 6.

<table>
<thead>
<tr>
<th>Range</th>
<th>Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>U+2100-U+214F Letterlike Symbols</td>
</tr>
<tr>
<td></td>
<td>U+2200-U+22FF Mathematical Operators</td>
</tr>
<tr>
<td></td>
<td>U+2250-U+225F Control Pictures</td>
</tr>
<tr>
<td></td>
<td>U+25A0-U+25FF Block Elements</td>
</tr>
<tr>
<td></td>
<td>U+2600-U+26FF Miscellaneous Symbols</td>
</tr>
<tr>
<td></td>
<td>U+2900-U+29FF Supplemental Arrows-B</td>
</tr>
<tr>
<td></td>
<td>U+2800-U+28FF Misc. Symbols and Arrows</td>
</tr>
</tbody>
</table>

Range 4^A 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.

```
\def\charsrange[4]{%n
  "\buffers\-10FF, "\adores\-10DF, "\buffers\-245F, "\adores\-28FF, "\adores\-2AFF, "\adores\-27FF, "\adores\-40DF, "\adores\-A95F, "\adores\-AB0F, "\adores\-FBFF, "\adores\-F8FF, "\adores\-FEFF, "\adores\-FEFF, "\adores\-1AFF, "\adores\-1AFF, "\adores\-1F0F, "\adores\-1F0F, \dots \text{ (characters in }2000\text{--}206F \text{ which are not in range 9)}
} % non-Japanese
```

Range 5^A Surrogates and Supplementary Private Use Areas.

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

Range 7^I Characters used in CJK languages, but not included in Adobe-Japan1-7. The block list is indicated in Table 7.
Table 7. Unicode blocks in predefined character range 7.

<table>
<thead>
<tr>
<th>Code Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>U+1100–U+11FF</td>
<td>Hangul Jamo</td>
</tr>
<tr>
<td>U+2FF8–U+2FFF</td>
<td>Ideographic Description Characters</td>
</tr>
<tr>
<td>U+3130–U+31BF</td>
<td>Hangul Compatibility Jamo</td>
</tr>
<tr>
<td>U+31C0–U+31EF</td>
<td>CJK Strokes</td>
</tr>
<tr>
<td>U+A000–U+A48F</td>
<td>Yi Syllables</td>
</tr>
<tr>
<td>U+A960–U+A97F</td>
<td>Hangul Jamo Extended-A</td>
</tr>
<tr>
<td>U+AC00–U+D7AF</td>
<td>Hangul Syllables</td>
</tr>
<tr>
<td>U+D7B0–U+D7FF</td>
<td>Hangul Jamo Extended-B</td>
</tr>
</tbody>
</table>

**Notes on U+0080–U+00FF**

You should treat characters in textttU+0080–U+00FF as **ALchar**, when you use traditional 8-bit fonts, such as the marvosym package.

For example, \texttt{\textbackslash{}frowny} which is provided by the marvosym package has the same codepoint as § (U+00A7). Hence, as previous versions of Lua\TeX\-ja, if these characters are treated as JAchars, then \texttt{\textbackslash{}frowny} produces " § " (in a Japanese font).

To avoid such situations, the default setting of Lua\TeX\-ja is changed in version 20150906.0 so that all characters U+0080–U+00FF are treated as **ALchar**.

If you want to output a character as **ALchar** and **JAchar** regardless the range setting, you can use \texttt{\ltjalchar} and \texttt{\ltjjachar} respectively, as the following example.

\begin{verbatim}
1 \gtfamily\large% default, ALchar, JAchar
2 ¶, \ltjalchar`¶, \ltjjachar`¶% default: ALchar
3 α, \ltjalchar`α, \ltjjachar`α% default: JAchar
4.2 kanjiskip and xkanjiskip

JAgIue is divided into the following three categories:

- Glues/kerns specified in JFM. If \texttt{\inhibitglue} is issued around a **JAchar**, this glue will not be inserted at the place.
- The default glue which inserted between two **JAchars** (**kanjiskip**).
- The default glue which inserted between a **JAchar** 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{\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 p\TeX.

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

\begin{verbatim}
1 kanjiskip: \ltjgetparameter{kanjiskip},\% kanjiskip: 0.0pt plus 0.4pt minus 0.5pt,
2 xkanjiskip: \ltjgetparameter{xkanjiskip}\% xkanjiskip: 2.40555pt plus 1.0pt minus 1.0pt

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 \texttt{kanjiskip} or \texttt{xkanjiskip} to \texttt{\maxdimen} (these "ideal width" cannot be retrived by \texttt{\ltjgetparameter}).

4.3 Insertion setting of xkanjiskip

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

15
The second argument of \texttt{\preonly} means that the insertion of \texttt{\kanjiskip} 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{\skankjiskip}, 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 \LaTeX, this is achieved by setting \texttt{\baselineskip} (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.

<table>
<thead>
<tr>
<th>Alphabetic fonts</th>
<th>\texttt{yalbaselineshift} parameter</th>
<th>\texttt{talbaselineshift} parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese fonts</td>
<td>\texttt{yjabaselineshift} parameter</td>
<td>\texttt{tjabaselineshift} parameter</td>
</tr>
</tbody>
</table>

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

\begin{verbatim}
\begin{verbatim}
\vrule width 150pt height 0.2pt depth 0.2pt \hskip-120pt
\end{verbatim}
\end{verbatim}
\begin{verbatim}
\ltjsetparameter{yjabaselineshift=0pt, yalbaselineshift=0pt}abcあいう
\end{verbatim}
\begin{verbatim}
\ltjsetparameter{yjabaselineshift=5pt, yalbaselineshift=2pt}abcあいう
\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 height 4.417pt depth -4.217pt\
\kern-150pt
\large 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, Lua\TeX-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 \texttt{\lpcode} and \texttt{\rpcode} are both non-zero.

This problem does not apply for \texttt{yjabaselineshift} nor \texttt{tjabaselineshift}, because a \texttt{\JAlchar} 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*),

\texttt{jxspmode, akxspmode, prebreakpenalty, postbreakpenalty} and \texttt{kcatcode}

are stored by each character codes.

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

\begin{verbatim}
\ltjsetparameter{postbreakpenalty={‘半’, 10}}
\ltjsetparameter{postbreakpenalty={‘半’, 20}}
\newcommand{showpostpena}[1]{% \leavevmode\setbox0=\hbox{#1}\hbox{}}% \unhbox0\setbox0=\lastbox\the\lastpenalty
\showpostpena{半}, \showpostpena{半}, \showpostpena{半}, \addjfontfeatures{CharacterWidth=Half}\showpostpena{半}}
\end{verbatim}
Part II

Reference

5 \texttt{\textbackslash catcode} in \LaTeXX-ja

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

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

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

5.2 Case of \LaTeXX-ja

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

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

Default setting of \texttt{\textbackslash catcode} of Unicode characters are located in \texttt{luatex-unicode-letters.tex}, which is based on \texttt{unicode-letters.tex} (for \TeX).

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

However, the default setting of \texttt{\textbackslash catcode} differs between \TeX\ and \LaTeXX, by the following reasons:

- (plain format) \texttt{luatex-unicode-letters.tex} is based on old \texttt{unicode-letters.tex}.
- The latter half of \texttt{unicode-letters.tex} and \texttt{unicode-letters.def} sets \texttt{\textbackslash catcode} of several characters to 11, via setting \texttt{\XeTeXcharclass}. However, this latter half does not exist (plain case), or not executed (\TeX case) in \LaTeXX.

In other words,

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

\textbf{\LaTeXX} In recent (2015-10-01 or later) \LaTeXX, 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 \TeX users to shifting to \LaTeXX-ja, since several control words containing Kanji or other fullwidth characters, such as \texttt{\textbackslash x} or \texttt{\textbackslash 1 年} are used in \TeX. Hence, \LaTeXX-ja have a counterpart of \texttt{unicode-letters.tex} for \LaTeXX, to match the \texttt{\textbackslash catcode} setting with that of \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 \TeX, in up\TeX, and in \LaTeXX-ja. Table 9 shows the difference. Except for four characters ",", "", ",", ",", \LaTeXX-ja admits more characters in a control word than 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}.\footnote{up\TeX divides U+FF00\textendash{}U+FFEF (Halfwidth and Fullwidth Forms) into three subblocks, and \texttt{\textbackslash catcode} can be set by a subblock.}
<table>
<thead>
<tr>
<th>\texttt{\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>Y</td>
<td>Y</td>
<td>ignored</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</th>
<th>col.</th>
<th>\texttt{p\TeX}</th>
<th>\texttt{up\TeX}</th>
<th>\texttt{Lua\TeX-ja}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(U+30FB)</td>
<td>1</td>
<td>6</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>(U+3098)</td>
<td>1</td>
<td>11</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>(U+309C)</td>
<td>1</td>
<td>12</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>(U+FF48)</td>
<td>1</td>
<td>14</td>
<td>N</td>
</tr>
<tr>
<td>5</td>
<td>(U+FF3E)</td>
<td>1</td>
<td>16</td>
<td>N</td>
</tr>
<tr>
<td>6</td>
<td>(U+FF3F)</td>
<td>1</td>
<td>18</td>
<td>N</td>
</tr>
<tr>
<td>7</td>
<td>(U+3003)</td>
<td>1</td>
<td>23</td>
<td>N</td>
</tr>
<tr>
<td>8</td>
<td>(U+4ED0)</td>
<td>1</td>
<td>24</td>
<td>N</td>
</tr>
<tr>
<td>9</td>
<td>(U+3005)</td>
<td>1</td>
<td>25</td>
<td>N</td>
</tr>
<tr>
<td>10</td>
<td>(U+3006)</td>
<td>1</td>
<td>26</td>
<td>N</td>
</tr>
<tr>
<td>11</td>
<td>(U+3007)</td>
<td>1</td>
<td>27</td>
<td>N</td>
</tr>
<tr>
<td>12</td>
<td>(U+30FC)</td>
<td>1</td>
<td>28</td>
<td>N</td>
</tr>
<tr>
<td>13</td>
<td>(U+FF8F)</td>
<td>1</td>
<td>31</td>
<td>N</td>
</tr>
<tr>
<td>14</td>
<td>(U+FF3C)</td>
<td>1</td>
<td>32</td>
<td>N</td>
</tr>
</tbody>
</table>

 Lua\TeX supports four \Omega-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.

 Lua\TeX-ja supports four directions, as shown in Table 10. The second column (\texttt{yoko} direction) is just horizontal writing, and the third column (\texttt{tate} direction) is vertical writing. The fourth column (\texttt{dtou} direction) is actually a hidden feature of p\TeX. We implemented this for debugging purpose. The fifth column (\texttt{utod} direction) corresponds the "\texttt{tate (math) direction}" of p\TeX.

Directions can be changed by \texttt{\yoko}, \texttt{\tate}, \texttt{\dtou}, \texttt{\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 \texttt{utod}, when the direction outside the math formula is \texttt{tate} (vertical writing).

6.1 Boxes in different direction

As in p\TeX, one can use boxes of different direction in one document. The below is an example.
Table 10. Directions supported by LuaTEX-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>Left</td>
<td>Right</td>
</tr>
<tr>
<td>Beginning of the line</td>
<td>yoko</td>
<td>tate</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>Used Japanese font</td>
<td>horizontal (\jfont)</td>
<td>vertical (\tfont)</td>
<td>horizontal (90° rotated)</td>
<td></td>
</tr>
</tbody>
</table>

Example

銀は、Ag

(Notation used in Ω)

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

\[\text{\textbf{\texttt{\textbackslash wd and direction}}\textbf{\textit{In \LaTeX}, \texttt{\textbackslash wd, \texttt{\textbackslash h}, \texttt{\textbackslash dp} means the dimensions of a box register \textit{with respect to the current direction}. This means that the value of \texttt{\textbackslash wd0 etc. might differ when the current direction is different, even if \texttt{\box0 stores the same box}. However, this no longer applies in Lua\LaTeX-ja.}}\]

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

\[\texttt{\textbackslash ltjgetwd(num), \texttt{\textbackslash ltjgetht(num)}, \texttt{\textbackslash ltjgetdp(num)}}\]

These commands return an \textit{internal dimension} of \texttt{\textbackslash box(num) with respect to the current direction.}

One can use these in \texttt{\dimexpr primitive, as the followings.}

\[\texttt{\dimexpr 2\texttt{\ltjgetwd42-3pt}\relax, \texttt{\the\ltjgetwd1701}}\]

The following is an example.

```
1 \parindent0pt
2 \setbox0=\hbox{\yokoよこぐみ}
3 \fboxsep=0mm\fbox{\copy0}
4 \vbox{\hsize=20mm\raggedleft}
   \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.}
5 \vbox{\hsize=20mm\raggedleft}
   \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.}
6 \vbox{\hsize=20mm\raggedleft}
   \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.}
7 \vbox{\hsize=20mm\raggedleft}
   \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.}
8 \vbox{\hsize=20mm\raggedleft}
   \the\ltjgetwd32767, \the\ltjgetht32767, \the\ltjgetdp32767.}
```

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

<table>
<thead>
<tr>
<th>typeset in yoko direction</th>
<th>typeset in tate or utod direction</th>
<th>typeset in dtou direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_y = h_y + d_y )</td>
<td>( W_t = h_t + d_t )</td>
<td>( W_D = h_D + d_D )</td>
</tr>
<tr>
<td>( H_y = w_y )</td>
<td>( H_t = w_t )</td>
<td>( H_D = w_D )</td>
</tr>
<tr>
<td>( D_y = 0 \text{ pt} )</td>
<td>( D_t = w_t/2 )</td>
<td>( D_D = h_D )</td>
</tr>
<tr>
<td></td>
<td>( D_T = h_T )</td>
<td></td>
</tr>
</tbody>
</table>

6.2 Getting current direction

The \texttt{direction} parameter returns the current direction, and the \texttt{boxdir} parameter (with the argument \( \langle \text{num} \rangle \)) returns the direction of a box register \texttt{\{box\}(\text{num})}. The returned value of these parameters are a \texttt{string}:

<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>

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

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

\texttt{\ltjsetwd42 \textbackslash 20pt, \ltjsetwd42=\textbackslash 20pt, \ltjsetwd=42 \textbackslash 20pt, \ltjsetwd=42=\textbackslash 20pt}

\texttt{\textbackslash留在mode\textbackslash def\textbackslash \textbackslash dir\textbackslash \textbackslash ltjgetparameter\textbackslash \{direction\}}

\texttt{\textbackslash hbox\textbackslash \{yoko\} \textbackslash \textbackslash dir, \textbackslash hbox\textbackslash \{tate\}\textbackslash \textbackslash dir,}
\texttt{\textbackslash hbox\textbackslash \{utod\}\textbackslash \textbackslash dir,}
\texttt{\textbackslash hbox\textbackslash \{utod\}\textbackslash \textbackslash dir,}
\texttt{\textbackslash hbox\textbackslash \{tate\}\textbackslash \textbackslash hbox\textbackslash \{tate\ math: \textbackslash dir\}4}
\texttt{\textbackslash \setbox2=\textbackslash hbox\textbackslash \{tate\}\textbackslash \ltjgetparameter\textbackslash \{boxdir\}\{2}
Table 12. Differences between horizontal JFMs shipped with \TeX-ja

<table>
<thead>
<tr>
<th>Blue: jfm-ujis.lua</th>
<th>Black: jfm-jis.lua</th>
<th>Red: jfm-min.lua</th>
</tr>
</thead>
<tbody>
<tr>
<td>ある日モモちゃんがお使いで迷子になって泣きました。</td>
<td>ちょっと！何</td>
<td>ちょっと!!何</td>
</tr>
<tr>
<td>漢つ</td>
<td>漢つ</td>
<td>漢つ</td>
</tr>
</tbody>
</table>

(Blue: jfm-ujis.lua, Black: jfm-jis.lua, Red: jfm-min.lua)

6.3 Overridden box primitives

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

\begin{itemize}
  \item \texttt{\unhbox\langle num\rangle}, \texttt{\unvbox\langle num\rangle}, \texttt{\unhcopy\langle num\rangle}, \texttt{\unvcopy\langle num\rangle}
  \item \texttt{\vadjust\langle material\rangle}
  \item \texttt{\insert\langle number\rangle\{\langle material\rangle\}}
  \item \texttt{\lastbox}
  \item \texttt{\raise\langle dimen\rangle\langle box\rangle}, \texttt{\lower\langle dimen\rangle\langle box\rangle} etc., \texttt{\vcenter}
  \item \texttt{\vcenter}
\end{itemize}

7 Font Metric and Japanese Font

7.1 \texttt{\jfont}

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

\begin{verbatim}
\jfont\tradgt={file:KozMinPr6N-Regular.otf:script=latn;%
  +trad;-kern;jfm=ujis} at 14pt
\end{verbatim}

Note that the defined control sequence (\texttt{\tradgt} in the example above) using \texttt{\jfont} is not a \texttt{font_def} token, but a macro. Hence the input like \texttt{\fontname\tradgt} causes an error. We denote control sequences which are defined in \texttt{\jfont} by \texttt{\jfont_cs}.

\begin{itemize}
  \item \textbf{JFM} a JFM has measurements of characters and glues/kerns that are automatically inserted for Japanese typesetting. The structure of JFM will be described in the next subsection. At the calling of \texttt{\jfont}, you must specify which JFM will be used for this font by the following keys:

\begin{verbatim}
\jfm=(\texttt{name})
\end{verbatim}

Specify the name of (horizontal) JFM. If specified JFM has not been loaded, Lua\TeX-ja search and load a file named \texttt{jfm-\langle name\rangle}.lua.

The following horizontal JFMs are shipped with Lua\TeX-ja:

\begin{itemize}
  \item \texttt{jfm-ujis.lua} A standard horizontal JFM in Lua\TeX-ja. This JFM is based on \texttt{upnmlminr-h.tfm}, a metric for UTF/OTF package that is used in up\TeX. When you use the \texttt{luatexja-otf} package, you should use this JFM.
\end{itemize}
Figure 1. Example of \texttt{jfmvar} key

```
ダイナミックダイクマ
ダイナミックダイクマ
ダイナミックダイクマ
ダイナミックダイクマ

ダイナミックダイクマ
ダイナミックダイクマ
ダイナミックダイクマ
ダイナミックダイクマ
```

Figure 2. Kerning information and \texttt{kanjiskip}

```
jfm-jis.lua A counterpart for jis.tfm, "JIS font metric" 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.

jfm-min.lua A counterpart for min10.tfm, which is one of the default Japanese font metric shipped with \TeX.

The difference among these three JFMs is shown in Table 12.

\texttt{jfmvar}=(\texttt{string}) Sometimes there is a need that ....

\textbf{Using kerning information in a font} Some fonts have information for inter-glyph spacing. Lua\TeX-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 2 for detail.

Note that in \texttt{\setmainfont} etc. which are provided by luatexja-fontspec package, kerning option is set \texttt{off} (\texttt{Kerning=Off}) by default, because of the compatibility with previous versions of Lua\TeX-ja.

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

\texttt{extend=(extend)} expand the font horizontally by \texttt{(extend)}.

23
slant=⟨slant⟩ slant the font.

Note that LuaTeX-ja doesn’t adjust JFMs by these extend and slant settings; you have to write new JFMs on purpose. For example, the following example uses the standard JFM jfm-ujis.lua, hence letter-spacing and the width of italic correction are not correct:

```latex
\jfont\E=KozMinPr6N-Regular:extend=1.5;jfm=ujis;-kern
E あいうえお
\jfont\S=KozMinPr6N-Regular:slant=1;jfm=ujis;-kern
S あいう
```

This functionality is disabled by -ltjksp specification.

```latex
\leavevmode
\ltjsetparameter{kanjiskip=0pt plus 3\zw}
\vrule\hbox to 15\zw{あ 「い」う , えお}\vrule
\jfont\G=file:KozMinPr6N-Regular.otf% jfm=ujisv\T=file:KozMinPr6N-Regular.otf:-vert% jfm=ujisv\U=file:ipaexm.ttf:jfm=ujisv
\vrule\hbox to 15\zw{あ 「い」う , えお}\vrule
```

### 7.2 \texttt{\font}

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

- OpenType Feature vrt2\footnote{If the font does not define vrt2 feature, use vert instead.} is automatically activated, unless vert and/or vrt2 are explicitly activated or deactivated (as the second line in the example below).

```latex
\texttt{\font}S=file:KozMinPr6N-Regular.otf% vrt2 is automatically activated
\texttt{\font}T=file:KozMinPr6N-Regular.otf:-vert% vert and vrt2 are not activated
\texttt{\font}U=file:ipaexm.ttf: jfm=ujisv
% vert is automatically activated, since this font does not have vrt2
```

- Sometimes vert and/or 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.

In this situation, LuaTeX-ja performs all replacements which is defined in vert feature for some scripts for some languages.

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

- One have to specify the name of vertical JFM in jfm=⟨name⟩. LuaTeX-ja ships following vertical JFMs:

  - \texttt{\jfont-ujisv.lua} A standard vertical JFM in LuaTeX-ja. This JFM is based on \texttt{upnm1mnr-v.tfm}, a metric for UTF/ORTF package that is used in \texttt{upTeX}.

  - \texttt{\jfont-tmin.lua} A counterpart for \texttt{tmin10.tfm}, which is one of the default Japanese font metric shipped with \texttt{upTeX}.

- If vert and/or 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 3).
7.3 Default Japanese fonts and JFMs

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

\ltj@stdmcfont The default Japanese font for the mincho family.
\ltj@stdgtfont The default Japanese font for the gothic family.
\ltj@stdyokojfm The default JFM for horizontal direction.
\ltj@stdtatejfm The default JFM for vertical direction.

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

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

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

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

7.4 Prefix psft

Besides “file” and “name” prefixes which are introduced in the luaotfload package, LuaTeX-ja adds “psft” prefix in \jfont (and \font), to specify a “name-only” Japanese font which will not be embedded to PDF. Note that these non-embedded fonts under current LuaTeX 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 “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 psft prefix, because they are only simple linear transformations.

---

*This is because commands has $@$ in their names.
The default font defined by using psft prefix is for Japanese typesetting; it is Adobe-Japan1-7 CID-keyed font. One can specify cid key to use other CID-keyed non-embedded fonts for Chinese or Korean typesetting.

\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

Note that the code above specifies 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.,
\jfont\test={psft:Ryumin-Light:cid=Adobe-Japan2;jfm=jis}

produces the following error:
!
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
\?

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 Lua\TeX-ja project team.
? 

7.5 Structure of a JFM file

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

version=(version) (optional, default value is 1)

The version JFM. Currently 1, 2, and, 3 are supported

dir=(direction) (required)

The direction of JFM. 'yoko' (horizontal) or 'tate' (vertical) are supported.

zw=(length) (required)

The amount of the length of the "full-width".

zh=(length) (required)

The amount of the "full-height" (height + depth).

kanjiskip={⟨natural⟩, ⟨stretch⟩, ⟨shrink⟩} (optional)

This field specifies the "ideal" amount of kanjiskip. As noted in Subsection 4.2, if the parameter
kanjiskip is \maxdimen, 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.

xkanjiskip={⟨natural⟩, ⟨stretch⟩, ⟨shrink⟩} (optional)

Like the kanjiskip field, this field specifies the "ideal" amount of xkanjiskip.
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 4. The position of the real glyph (horizontal Japanese fonts)

■ Character classes

Besides from above fields, a JFM file have several sub-tables those indices are natural numbers. The table indexed by \( i \in \omega \) stores information of character class \( i \). At least, the character class 0 is always present, so each JFM file must have a sub-table whose index is [\( 0 \)]. Each sub-table (its numerical index is denoted by \( i \)) has the following fields:

- \( \text{chars} = \{ (\text{character}), \ldots \} \) (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:
  - a Unicode code point
  - the character itself (as a Lua string, like ‘あ’)
  - a string like ‘あ*’ (the character followed by an asterisk)
  - several "imaginary" characters (We will describe these later.)

- \( \text{width} = \langle \text{length} \rangle \), \( \text{height} = \langle \text{length} \rangle \), \( \text{depth} = \langle \text{length} \rangle \), \( \text{italic} = \langle \text{length} \rangle \) (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.

- \( \text{left} = \langle \text{length} \rangle \), \( \text{down} = \langle \text{length} \rangle \), \( \text{align} = \langle \text{align} \rangle \)
  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 4 and Figure 5.
  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’.

- \( \text{kern} = \{ [j] = \langle \text{kern} \rangle, \ldots \} \)
- \( \text{glue} = \{ [j] = \langle \text{width} \rangle, \langle \text{stretch} \rangle, \langle \text{shrink} \rangle, \ldots \} \)
  Specifies the amount of kern or glue which will be inserted between characters in character class \( i \) and those in character class \( j \).
\[\text{Figure 5. The position of the real glyph (vertical Japanese fonts)}\]

\(\langle \text{ratio} \rangle\) 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 \(\langle \text{ratio} \rangle\) to \(0.25/(0.5 + 0.25) = 1/3\).

In case of glue, one can specify following additional keys in each \(\langle j \rangle\) subtable:

**priority=\(\langle \text{priority} \rangle\)** An integer in \([-4, 3]\) (treated as 0 if omitted), or a pair of these integers \{\(\langle \text{stretch} \rangle\), \(\langle \text{shrink} \rangle\)\} (version 2 or later). This is used only in line adjustment with priority by luatexja-adjust (see Subsection 11.3). Higher value means the glue is easy to stretch, and is also easy to shrink.

**kanjiskip\_natural=\(\langle \text{num} \rangle\), kanjiskip\_stretch=\(\langle \text{num} \rangle\), kanjiskip\_shrink=\(\langle \text{num} \rangle\)**

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 kanjiskip times the value of 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 kanjiskip times the value of 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 kanjiskip.

Hence we have the following result:

```latex
\leavevmode
\ltjsetparameter{kanjiskip=0pt plus 3zw}
\vrule\box to 15zw{あ「い」う, えお}\vrule
\vrule\box{あ「い」う, えお}\vrule\par
\ltjsetparameter{kanjiskip=0pt minus \zw}
\vrule\box to 6.5zw{あ「い」う, えお}\vrule
```

**end\_stretch=\(\langle \text{kern} \rangle\), end\_shrink=\(\langle \text{kern} \rangle\)** (optional, version 1 only)

**end\_adjust=\{\(\langle \text{kern} \rangle\), \(\langle \text{kern} \rangle\), ...\}** (optional, version 2 or later)

**Character to character classes** We explain how the character class of a character is determined, using jfm-test.lua which contains the following:
Now consider the following input/output:

```
1. \jfont\a=file:KozMinPr6N-Regular.otf:jfm=test;+hwid
2. \setbox0\hbox{\a \ヒ 漢}
3. \the\wd0 15.0pt
```

Now we look why the above source outputs 15 pt.

1. The character “ヒ” is converted to its half width form “ひ” by \hwid feature.
2. According to the JFM, the character class of “ひ” 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 \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, LuaTeX-ja adopts the class determined by the glyph before application of features. The following input is an example.

```
1. \jfont\a=file:KozMinPr6N-Regular.otf:jfm=test;+vert
2. \a 漢。\inhibitglue 漢 漢
```

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

1. As the case of “ひ”, the ideographic full stop “。” is converted to its vertical form “тир” (U+FE12) by \vert feature.
2. The character class of “тир”, according to the JFM is zero.
3. However, LuaTeX-ja remembers that this “тир” 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, you can specify several “imaginary characters” in \chars field. The most of these characters are regarded as the characters of class 0 in \ptex. As a result, LuaTeX-ja can control typesetting finer than \ptex. 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 \noindent) paragraph. If a hbox \b begins (resp. ends) a glue or kern between this “charater” and a \jaChar, \jaGlue won’t be inserted before(resp. after) the hbox \b. \kanjiskip and \xkanjiskip around a hbox.

- **'parbdd'**
  The beginning of an (indented) paragraph.

- **'jcharbdd'**
  A boundary between \jaChar and anything else.

- **'alchar', 'nox_alchar'**
  (version 3 or later) A boundary between \jaChar and \alChar.

- **'glue'**
  (version 3 or later) A boundary between \jaChar, and, a glue or kern.

-1 The left/right boundary of an inline math formula.
Table 14. Commands for Japanese math fonts

<table>
<thead>
<tr>
<th>Japanese fonts</th>
<th>alphabetic fonts</th>
</tr>
</thead>
<tbody>
<tr>
<td>\jfam ∈ [0, 256)</td>
<td>\fam</td>
</tr>
<tr>
<td>jatextfont=((jfam),(jfont_cs))</td>
<td>textfont((fam)=(font_cs))</td>
</tr>
<tr>
<td>jascriptfont=((jfam),(jfont_cs))</td>
<td>scriptfont((fam)=(font_cs))</td>
</tr>
<tr>
<td>jascriptscriptfont=((jfam),(jfont_cs))</td>
<td>scriptscriptfont((fam)=(font_cs))</td>
</tr>
</tbody>
</table>

Porting JFM from \PLTEX

See Japanese version of this manual.

7.6 Math font family

\PLTEX\ handles fonts in math formulas by 16 font families\(^\text{10}\), 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 \PLTEX\’s primitives for math font families. There is no relation between the value of \fam and that of \jfam; with appropriate settings, you can set both \fam and \jfam to the same value. Here (jf_{font cs}) in the argument of jatextfont etc. is a control sequence which is defined by \jfont, i.e., a horizontal Japanese font.

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.

\texttt{luatexja.load_jfm} callback

With this callback you can overwrite JFMs. This callback is called when a new JFM is loaded.

\begin{verbatim}
function (〈table〉 jfm_info, 〈string〉 jfm_name)
    return 〈table〉 new_jfm_info
end
\end{verbatim}

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

An example of this callback is the \texttt{ltjarticle} class, with forcefully assigning character class 0 to 'parbdd' in the JFM jfm-min.lua.

\texttt{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 \texttt{luatexja.define_font} callback is called just when new Japanese font is loaded.

\begin{verbatim}
function (〈table〉 jfont_info, 〈number〉 font_number)
    return 〈table〉 new_jfont_info
end
\end{verbatim}

\texttt{jfont_info} has the following fields, which may not overwritten by a user:

- \texttt{size} The font size specified at \texttt{jfont} in scaled points (1 sp = 2\(^{-16}\) pt).
- \texttt{zw}, \texttt{zh}, \texttt{kanjiskip}, \texttt{xkanjiskip} These are scaled value of those specified by the JFM, by the font size.
- \texttt{jfm} The internal number of the JFM.
- \texttt{var} The value of jf\_var key, which is specified at \texttt{jfont}. The default value is the empty string.
- \texttt{chars} The mapping table from character codes to its character classes.

The specification \{i\}.chars={\langle character\rangle, ...} in the JFM will be stored in this field as \texttt{chars}={{\langle character\rangle}=\texttt{i}, ...}.

\(^{10}\text{Omega, Aleph, \PLTEX} and \texttt{up\PLTEX} can handles 256 families, but an external package is needed to support this in plain \PLTEX\ and \LUATEX.\)
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:

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

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-off 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:

```lua
1 function (number char_class, table jfont_info, number chr_code)
2     if char_class~=0 then return char_class
3     else
4         ....
5         return (number new_char_class or 0)
6     end
7     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.

luatexja.set_width callback

This callback is called when LuaTEX-ja is trying to encapsule a JAchar glyph_node, to adjust its dimension and position.

```lua
1 function (table shift_info, table jfont_info, table char_type)
2     return (table new_shift_info)
3 end
```

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 \ltjsetparameter

As described before, \ltjsetparameter and \ltjgetparameter are commands for accessing most parameters of LuaTeX-ja. One of the main reasons that LuaTeX-ja didn’t adopt the syntax similar to that of p\TeX (e.g., \prebreakpenalty = 10000) is the position of hpack_filter callback in the source of LuaTeX, see Section 12.

\ltjsetparameter and \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; \ltjsetparameter does a local assignment and \ltjglobalsetparameter does a global one by default. They also obey the value of globaldefs, like other assignments.

The following is the list of parameters which can be specified by the \ltjsetparameter command. \[cs\] indicates the counterpart in p\TeX, 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.

\textbf{jcharwidowpenalty}=(\texttt{\*}) ∗ [jcharwidowpenalty]

Penalty value for suppressing orphans. This penalty is inserted just after the last JAchar which is not regarded as a (Japanese) punctuation mark.

\textbf{kcatcode}=(⟨chr_code⟩,⟨natural number⟩)∗

An additional attributes which each character whose character code is (\texttt{chr_code}) has. At version 20120506.0 or later, the lowermost bit of (\texttt{natural number}) indicates whether the character is considered as a punctuation mark (see the description of \texttt{jcharwidowpenalty} above).

\textbf{prebreakpenalty}=(⟨chr_code⟩,⟨penalty⟩)∗ [prebreakpenalty]

Set a penalty which is inserted automatically before the character (\texttt{chr_code}), to prevent a line starts from this character. For example, a line cannot started with one of closing brackets ” ) “, so Lua\TeX-ja sets
\begin{verbatim}
\ltjsetparameter{prebreakpenalty={`) ,10000}}
\end{verbatim}

by default.

p\TeX has following restrictions on prebreakpenalty and postbreakpenalty, but they don’t exist in Lua\TeX-ja:

- Both prebreakpenalty and postbreakpenalty cannot be set for the same character.
- We can set prebreakpenalty and postbreakpenalty up to 256 characters.

\textbf{postbreakpenalty}=(⟨chr_code⟩,⟨penalty⟩)∗ [postbreakpenalty]

Set a penalty which is inserted automatically after the character (\texttt{chr_code}), to prevent a line ends with this character.

\textbf{jatextfont}=(⟨jfam⟩,⟨jfont_cs⟩)∗ [textfont in \TeX]

\textbf{jascriptfont}=(⟨jfam⟩,⟨jfont_cs⟩)∗ [scriptfont in \TeX]

\textbf{jascriptscriptfont}=(⟨jfam⟩,⟨jfont_cs⟩)∗ [scriptscriptfont in \TeX]

\textbf{yjabaselineshift}=(\texttt{dimen})
\textbf{yalbaselineshift}=(\texttt{dimen}) [ybaselineshift]
\textbf{tjabaselineshift}=(\texttt{dimen})
\textbf{talbaselineshift}=(\texttt{dimen}) [tbaselineshift]
\texttt{jaxspmode} = \{(\texttt{chr\_code}), (\texttt{mode})\}^* \\
Set whether inserting \texttt{xkanjiskip} is allowed before/after a \texttt{JAchar} whose character code is \texttt{chr\_code}. 
The followings are allowed for \texttt{(mode)}:

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

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

\texttt{alxspmode} = \{(\texttt{chr\_code}), (\texttt{mode})\}^* \texttt{[\xspcode]}
Set whether inserting \texttt{xkanjiskip} is allowed before/after a \texttt{ALchar} whose character code is \texttt{chr\_code}. 
The followings are allowed for \texttt{(mode)}:

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

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

\texttt{autospacing} = \langle \texttt{bool} \rangle \texttt{[\autospacing]}
\texttt{autoxspacing} = \langle \texttt{bool} \rangle \texttt{[\autoxspacing]}
\texttt{kanjiskip} = \langle \texttt{skip} \rangle^* \texttt{[\kanjiskip]}
The default glue which inserted between two \texttt{JAchars}. Changing current Japanese font does not alter this parameter, as \LaTeX.
If the natural width of this parameter is \texttt{\maxdimen}, Lua\TeX-ja uses the value which is specified in the JFM for current Japanese font (See Subsection 7.5).

\texttt{xkanjiskip} = \langle \texttt{skip} \rangle^* \texttt{[\xkanjiskip]}
The default glue which inserted between a \texttt{JAchar} and an \texttt{ALchar}. Changing current font does not alter this parameter, as \LaTeX.
As \texttt{kanjiskip}, if the natural width of this parameter is \texttt{\maxdimen}, Lua\TeX-ja uses the value which is specified in the JFM for current Japanese font (See Subsection 7.5).

\texttt{differentjfm} = \langle \texttt{mode} \rangle^*
Specify how glues/kerns between two \texttt{JAchars} whose JFM (or size) are different. The allowed arguments are the followings:

average, both, large, small, pleft, pright, paverage

The default value is \texttt{paverage}. ...

\texttt{jacharrange} = \langle \texttt{ranges} \rangle
\texttt{kansujichar} = \{(\texttt{digit}), (\texttt{chr\_code})\}^* \texttt{[\kansujichar]}
\texttt{direction} = \langle \texttt{dir} \rangle \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 \ltjgetparameter

\ltjgetparameter is a control sequence for acquiring parameters. It always takes a parameter name as first argument.

\ltjgetparameter{differentjfm}, \ltjgetparameter{autospacing}, \ltjgetparameter{kanjiskip}, \ltjgetparameter{prebreakpenalty}\{\`\}.

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

• If first argument is one of the following, no additional argument is needed.
  jcharwidowpenalty, yjabaselineshift, yalbaselineshift, autospacing, autoxspacing, kanjiskip, xkanjiskip, differentjfm, direction

Note that \ltjgetparameter{autospacing} and \ltjgetparameter{autoxspacing} returns 1 or 0, not true nor false.

• If first argument is one of the following, an additional argument—a character code, for example—is needed.
  kcatcode, prebreakpenalty, postbreakpenalty, jaxspmode, alxspmode

\ltjgetparameter{jaxspmode}\{\ldots\} and \ltjgetparameter{alxspmode}\{\ldots\} returns 0, 1, 2, or 3, instead of preonly etc.

• For an integer \langle digit \rangle between 0 and 9, \ltjgetparameter{kansujichar}\{\langle digit \rangle\} returns the character code of the result of \kansuji\langle digit \rangle.

• \ltjgetparameter{adjustdir} returns a integer which represents the direction of the surrounding vertical list. As direction, the return value 1 means down-to-up direction, 3 means tate direction (vertical typesetting), and 4 means yoko direction (horizontal typesetting).

• For an integer \langle reg_num \rangle between 0 and 65535, \ltjgetparameter{boxdim}\{\langle reg_num \rangle\} returns the direction of \box\langle reg_num \rangle. If this box register is void, the returned value is zero.

• The following parameter names cannot be specified in \ltjgetparameter.

  jatextfont, jascriptfont, javascriptscriptfont, jacharrange

• \ltjgetparameter{chartorange}\{\langle chr_code \rangle\} returns the range number which \langle chr_code \rangle belongs to (although there is no parameter named "chartorange").

  If \langle chr_code \rangle is between 0 and 127, this \langle chr_code \rangle does not belong to any character range. In this case, \ltjgetparameter{chartorange}\{\langle chr_code \rangle\} returns −1.

Hence, one can know whether \langle chr_code \rangle is JChar or not by the following:

\ltjgetparameter{jacharrange}\{\ltjgetparameter{chartorange}\{\langle chr_code \rangle\}\}\% 0 if JChar, 1 if ALChar

• Because the returned value is string, the following conditionals do not work if kanjiskip (or xkanjiskip) has the stretch part or the shrink part.

  \ifdim\ltjgetparameter{kanjiskip}\textasciitilde\textasciitilde \fi
  \ifdim\ltjgetparameter{xkanjiskip}\textasciitilde\textasciitilde \fi

  The correct way is using a temporary register.

  \@tempskipa=\ltjgetparameter{kanjiskip} \ifdim\@tempskipa\textasciitilde\textasciitilde \fi
  \@tempskipa=\ltjgetparameter{xkanjiskip} \ifdim\@tempskipa\textasciitilde\textasciitilde \fi

  paverage, 1, 0.0pt plus 0.4pt minus 0.5pt, 10000.
8.3 Alternative Commands to \ltjsetparameter

The basic method to set parameters of LuaTeX-ja is to use \ltjsetparameter or \ltjglobalsetparameter. However, these commands are slow, because they parse a key-value list, so several alternative commands are used in LuaTeX-ja. *This subsection is not for general LuaTeX-ja users.*

■ Setting \texttt{kanjiskip} or \texttt{xkanjiskip}  
In \texttt{ltjsclasses}, every size-changing command such as \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 \ltjsetparameter:

\begin{verbatim}
\ltjsetparameter{kanjiskip=0\zw plus .1\zw minus .01\zw}
\@tempskipa=\ltjgetparameter{xkanjiskip}
\ifdim\@tempskipa>\z@
  \if@slide
    \ltjsetparameter{xkanjiskip=0.1em}
  \else
    \ltjsetparameter{xkanjiskip=0.25em plus 0.15em minus 0.06em}
  \fi
\fi
\end{verbatim}

Hence, LuaTeX-ja defines more primitive commands, namely \texttt{\ltjsetpar\global}, \texttt{\ltjsetkanjiskip}, and \texttt{\ltjsetxkanjiskip}. Here \texttt{\ltjsetpar\global\ltjsetkanjiskip 10pt}
and \texttt{\ltjsetparameter{kanjiskip=10pt}} has the same effect. The actual code of \texttt{ltjsclasses} is shown below:

\begin{verbatim}
\ltjsetparameter{kanjiskip=0\zw 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 \texttt{\ltjsetkanjiskip} or \texttt{\ltjsetxkanjiskip} alone, that is, without executing \texttt{\ltjsetpar\global} in advance, is not supported.

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

9.1 Commands for compatibility with \TeX{}

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

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

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

\begin{verbatim}
1 \newcount\hoge
2 \hoge="2423 %"
3 \the\hoge, \kansuji\hoge\}
4 \jis\hoge, \char\jis\hoge\}
5 \kansuji1701
\end{verbatim}

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

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

35
9.2 \inhibitglue

\inhibitglue suppresses the insertion of JAglue. 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=file:KozMinPr6N-Regular.otf:jfm=test \g
\fbox{\hbox{あウあ\inhibitglue ウ}}
\inhibitglue\par\noindent あ1
\par\noindent\inhibitglue\par\noindent あ2
\par\hrule
\noindent\inhibitglue あ office
\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.
- 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.

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.

\begin{verbatim}
\parindent1\zw
\noindent ああああああああ\par % for comparison
「あああああ\par % normal paragraph
\everypar{\null}\ltjfakeparbegin
「あああああ
\end{verbatim}

\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:

\begin{verbatim}
\parindent1\zw
\noindent ああああああああ\par % for comparison
「あああああ\par % normal paragraph
\everypar{\null}\ltjfakeparbegin
「あああああ
\end{verbatim}

9.4 \ltjdeclarealtfont

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

\begin{verbatim}
\ltjdeclarealtfont⟨base_font_cs⟩⟨alt_font_cs⟩{(range)}
\end{verbatim}

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: \( -n \) \((n \geq 1)\) means that all characters of character classes \( n \), with respect to JFM used by \langle base_font_cs \rangle. Note that characters which do not exist in \langle alt_font_cs \rangle are ignored.

For example, if \texttt{\hoge} uses \texttt{fjm-ujis.lua}, the standard JFM of Lua\TeX\-ja, then
\begin{verbatim}
\ltdeclarealtfont\hoge\piyo{"3000--30FF, {-1}--{-1}}
\end{verbatim}
does

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

Note that specifying negative numbers needs specification like \{-1\}--\{-1\}, because simple \texttt{\texttt{-1}} is treated as the range between 0 and 1.

10 Commands for \LaTeX\2\e

10.1 Loading Japanese fonts in \LaTeX\2\e

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]).

- \texttt{\selectfont} loads (and chooses) only the Japanese font for 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 (\texttt{\yoko}, \texttt{\tate}, \texttt{\dtou}, \texttt{\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 \texttt{\ltj@@orig@yoko} etc.

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

\begin{verbatim}
% in horizontal direction \(\yoko\)
\DeclareFixedFont\Hoge\{JT3\}\{gt\}\{m\}\{n\}\{12\} % JT3: for vertical direction
\Hoge
\setbox0=\hbox{\tate あいう}
\end{verbatim}

10.2 Patch for NFSS2

Japanese patch for NFSS2 in Lua\TeX\-ja is based on \texttt{plfonts.dtx} which plays the same role in \texttt{plfonts2.e}. We will describe commands which are not described in Subsection 3.1.

additional dimensions

Like \texttt{plfonts2.e}, Lua\TeX\-ja defines the following dimensions for information of current Japanese font:

\begin{verbatim}
\cht (height), \cdp (depth), \cHT (sum of former two),
\cwd (width), \cvs (lineskip), \chs (equals to \cwd)
\end{verbatim}

and its \texttt{\normalsize} version:
\Chd (height), \Ldp (depth), \Cwd (width), 
\Cvs (equals to \baselineskip), \Chs (equals to \cwd).

Note that \cwd and \chf 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))
\DeclareHiraganaKanjiEncoding((encoding))\{(text-settings)\}((math-settings))
\DeclareKatakanaKanjiEncoding((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)\{(shape)\})

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)\}

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

This command is almost same as \DeclareRelationFont, except that this command does a local
assignment, where \DeclareRelationFont does a global assignment.

\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:

\makeatletter
\SetRelationFont(JY3){\k@family}{m}{n}{TU}{lmss}{m}{n}
% \k@family: current Japanese font family
\userelfont\selectfont あいう abc
\makeatother

\adjustbaseline

In p\PDF\TeX 2e, \adjustbaseline sets \baselineskip to match the vertical center of “M” and that of
“\hbox{a}” in vertical typesetting:

\baselineskip \leftarrow \frac{(h_M + d_M) - (h_a + d_a)}{2} + d_M - d_a, 

where \(h_M\) and \(d_a\) denote the height of “\(\)" and the depth, respectively. In Lua\TeX-ja, this
\adjustbaseline does similar task, namely setting the \emph{talbaselineshift} parameter (a Japanese charac-
ter whose character class is zero is used, instead of ‘\hbox{a}’).

\family{(family)}

As in \PDF\TeX 2e, this command changes current font family (alphabetic, Japanese, or both) to \(\family\).
See Subsection 10.3 for detail.

\DeclareAlternateKanjiFont((base-encoding)\{(base-family)\}((base-series)\{(base-shape)\}
\{(alt-encoding)\{(alt-family)\}((alt-series)\{(alt-shape)\}(range)\}

As \texttt{\textbackslash ltjdeclarealtfont} (Subsection 9.4), characters in \(\family\) of the Japanese font (we say the \textit{base font}) which specified by first 4 arguments are typeset by the Japanese font which specified by fifth
to eighth arguments (we say the \textit{alternate font}). An example is shown in \textbf{Figure 6}. 38
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 6. An example of \DeclareAlternateKanjiFont

• 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.

・…

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:

\DeclareSymbolFont{mincho}{JY3}{mc}{m}{n}
\DeclareSymbolFontAlphabet{\mathmc}{mincho}

10.3 Detail of \fontfamily command

In this subsection, we describe when \fontfamily{family} changes current Japanese/alphabetic font family. Basically, current Japanese font family is changed to \textlangle family \textrangle 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 \textlangle family \textrangle, and another case that \textlangle family \textrangle 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 \LaTeX, but we re-implemented in Lua. We use an auxiliary list \textlangle \textlangle family \textrangle \textrangle. \\

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

2. If the family \textlangle family \textrangle has been listed in a list \textlangle \textlangle family \textrangle \textrangle, this means that \textlangle family \textrangle is not a Japanese font family.

3. If the luatexja-fontspec package is loaded, we stop here, and \textlangle family \textrangle 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 \textlangle enc \textrangle such that a font definition named \textlangle enc \textlangle family \textrangle \textrangle.fd (the file name is all lowercase) exists. If so, \textlangle family \textrangle is recognized as a Japanese font family (the font definition file won’t be loaded here). If not, \textlangle family \textrangle is not a Japanese font family, and \textlangle family \textrangle is appended to the list \textlangle \textlangle family \textrangle \textrangle.

■ Recognition as alphabetic font family

Next, whether alphabetic font family will be changed is determined in following order. We use auxiliary lists \textlangle \textlangle family \textrangle \textrangle and \textlangle \textlangle family \textrangle \textrangle.

1. If the family \textlangle family \textrangle has been listed in a list \textlangle \textlangle family \textrangle \textrangle, \textlangle family \textrangle is recognized as an alphabetic font family.
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>

2. If the family ⟨family⟩ has been listed in a list \( N_\alpha \), this means that ⟨family⟩ is not an alphabetic font family.

3. If there exists an alphabetic font encoding such that the family ⟨family⟩ has been defined under it, ⟨family⟩ is recognized as an alphabetic font family, and to memorize this, ⟨family⟩ is appended to the list \( F_\alpha \).

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

Also, each call of \( \textbackslash \DeclareFontFamily \) after loading of LuaTeX-ja makes the second argument ⟨family⟩ is appended to the list \( F_\alpha \).

The above order is very similar to \( \textbackslash \textsf{fontfamily} \) in \LaTeX, but more complicated (clause 3.). This is because \LaTeXe is a format however LuaTeX-ja is not, hence LuaTeX-ja does not know calls of \( \textbackslash \texttt{DeclareFontFamily before itself is loaded.} \)

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 \( \textbackslash \texttt{DeclareTextSymbol} \)

From \LaTeXe 2017/01/01, the standard encoding of LuaTeX is changed to the \( \texttt{TU} \) 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 \( \textbackslash \texttt{DeclareTextSymbol} \), and reloads \texttt{tuenc.def}.

Under original definition of \( \textbackslash \texttt{DeclareTextSymbol} \), internal commands which is defined by \( \textbackslash \texttt{DeclareTextSymbol} \) (such as \( \texttt{T1\textbackslash\textquotedblleft} \)) are \texttt{chardef} tokens. However, this no longer holds in LuaTeX-ja; for example, the meaning of \( \texttt{TU\textbackslash\textquotedblleft} \) is \texttt{\textbackslash itj\textbackslash\textbackslash\char8220}.

10.5 \strutbox

As \LaTeXe (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 Addon packages

LuaTeX-ja has several addon packages. These addons are written as \LaTeX packages, but luatexja-otf and luatexja-adjust can be loaded in plain LuaTeX by \textbackslash input.

11.1 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 fonts, the following “font features” specifications are allowed for the commands of Japanese version:

\begin{verbatim}
\jfonts[1]
  \ YokoFeatures={\{FakeBold=3\}}, TateFeatures={\{FakeBold=6\}},
  \ TateFont=KozGoPr6N-Regular
\jfonts[2]
\hbox{\yoko 横組のテスト}\hbox{\tate 縦組のテスト}
\addjfontfeatures{FakeBold=0}
\hbox{\yoko 横組}\hbox{\tate 縦組}
\end{verbatim}

Figure 7. An example of TateFeatures etc.

\begin{verbatim}
\jfonts[3]
  \ AltFont=
    \{}{\{\{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\ \{\
\end{verbatim}

Figure 8. An example of AltFont

Japanese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、わが国全土にわたつて自由のもたらす恵沢を確保し、……

Kanjiskip=(bool)

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 demonstration is shown in Figure 7.

YokoFeatures={⟨features⟩}

The YokoFeatures key specifies font features which are only turned on in horizontal writing. A demonstration is shown in Figure 7.

AltFont

As \ltjdeclarealtfont (Subsection 9.4) 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 demonstration is shown in Figure 8.

- **Remark on AltFont, YokoFeatures, TateFeatures keys**
  In AltFont, YokoFeatures, TateFeatures keys, one cannot specify per-shape settings such as BoldFeatures. For example,
  
  ```latex
  AltFont = {
    { Font=HogeraMin-Light, BoldFont=HogeraMin-Bold, Range="3000-“30FF, BoldFeatures={FakeBold=1} } }
  
  ```
  
  does not work. Instead, one have to write
  
  ```latex
  UprightFeatures = {
    AltFont = { { Font=HogeraMin-Light, Range="3000-“30FF, } },
    BoldFeatures = { 
      AltFont = { { Font=HogeraMin-Bold, Range="3000-“30FF, FakeBold=1 } },
  
  
  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.
  
  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 7 has no effect, because a color specification is already done in YokoFeatures and TateFeatures keys.

### 11.2 luatexja-otf

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

\begin{itemize}
  \item \texttt{\CID{⟨number⟩}}
    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, LuaTeX-ja treats that \texttt{⟨number⟩} is a CID number of Adobe-Japan1 character collection, and tries to typeset a "most suitable glyph".
  
  \item \texttt{\UTF{⟨hex number⟩}}
    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.
\end{itemize}

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

- **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{JA\char}s.
    \item In vertical direction, \texttt{\ver/\vrt2} 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 luaotfload package is not performed to these characters.
  \end{itemize}

\footnote{Useful macros by iNOUE Koichi, for the japanese-otf package.}
以上の原理は、「包除原理」とよく呼ばれるが

Additional syntax of JFM

The package luatexja-otf extends the syntax of JFM; the entries of \char's table in JFM now allows a string in the form 'AJ1-xxx', which stands for the character whose CID number in Adobe-Japan1 is xxx.

This extended notation is used in the standard JFM jfm-ujis.lua to typeset halfwidth Hiragana glyphs (CID 516–598) in halfwidth.

11.3 luatexja-adjust

(see Japanese version of this manual)

11.4 luatexja-ruby

This addon package provides functionality of “ruby” (furigana) annotations using callbacks of LuaTEx-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.

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:

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

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

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.
11.5 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 \texttt{<dir>}, which specifies the direction, as follows:

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

\begin{itemize}
  \item \texttt{y} \textit{yoko} direction (horizontal writing)
  \item \texttt{t} \textit{tate} direction (vertical writing)
  \item \texttt{z} \textit{utod} direction if direction outside the env. is \textit{tate}.
  \item \texttt{d} \textit{dtou} direction
  \item \texttt{u} \textit{utod} direction
\end{itemize}

\texttt{\parbox}<\texttt{dir}>\{⟨pos⟩\}{⟨width⟩}{⟨contents⟩}

This command typeset ⟨contents⟩ in LR-mode, in ⟨dir⟩ direction. If ⟨width⟩ is positive, the width of the box becomes this ⟨width⟩. In this case, ⟨contents⟩ will be aligned to left (when ⟨pos⟩ is 1), center (c), or right (r).

\texttt{\begin{picture}<dir>(x_size, y_size)(x_offset,y_offset) \ldots \end{picture}}

\texttt{\rensuj\{⟨pos⟩\}{⟨contents⟩}}
\texttt{\rensuj\{⟨counter_name⟩\}}
\texttt{\Kanji\{⟨counter_name⟩\}}
\texttt{\kasan\{⟨contents⟩\}}
\texttt{\bou\{⟨contents⟩\}}
\texttt{\boutenchar}

参照番号

11.6 luatexja-preset

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

Options which are given in \texttt{\usepackage} but not described in this subsection are simply passed to the \texttt{luatexja-fontspec} option. For example, the line 5 in below example is equivalent to lines 1–3.

\texttt{\usepackage[no-math]{fontspec}}
\texttt{\usepackage[match]{luatexja-fontspec}}
\texttt{\usepackage[kozuka-pr6n]{luatexja-preset}}
\texttt{\usepackage[no-math,match,kozuka-pr6n]{luatexja-preset}}
\texttt{\usepackage[no-math,match,kozuka-pr6n]{luatexja-preset}}

\textsubscript{12}If \texttt{nfssonly} option is not specified; in this case these options are simply ignored.
11.6.1 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}

nfssonly
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 packagefontspec 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.

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

nodeluxe (enabled by default)
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.

deluxe
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. 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.

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

bold
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.

jis90, 90jis
Use JIS X 0208:1990 glyph variants if possible.

jis2004, 2004jis
Use JIS X 0213:2004 glyph variants if possible.

jfm_yoko=(jfm)
Use the JFM jfm-(jfm).lua for horizontal direction, instead of jfm-ujis.lua (default JFM).

jfm_tate=(jfm)
Use the JFM jfm-(jfm).lua for vertical direction, instead of jfm-ujisv.lua (default JFM).

13Provided by \mgfamily and \texttag, because “rounded gothic” is called maru gothic (丸ゴシック) in Japanese.
jis Same as jfm.yoko=jis.

Note that jis98, 98jis, jis2004 and 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.

11.6.2 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** Kozuka Pro (Adobe-Japan1-4) fonts.
**kozuka-pr6** Kozuka Pr6 (Adobe-Japan1-6) fonts.
**kozuka-pr6n** Kozuka Pr6N (Adobe-Japan1-6, JIS04-savvy) fonts.

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>mincho</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>gothic</td>
<td>medium</td>
<td>KozGoPro-Regular*</td>
<td>KozGoProVI-Regular*</td>
<td>KozGoPr6N-Regular*</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>
</tbody>
</table>

**hiragino-pro** Hiragino Pro (Adobe-Japan1-5) fonts.
**hiragino-pron** Hiragino ProN (Adobe-Japan1-5, JIS04-savvy) fonts.

Hiragino fonts (except Hiragino Mincho W2) are bundled with Mac OS X 10.5 or later. Some editions of a Japanese word-processor “一太郎 2012” includes Hiragino ProN fonts. 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>mincho</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>gothic</td>
<td>medium</td>
<td>Hiragino Kaku Gothic Pro W3*</td>
<td>Hiragino Kaku Gothic ProN W3*</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 Std W8</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>Hiragino Maru Gothic Pro W4</td>
<td>Hiragino Maru Gothic ProN W4</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>family</th>
<th>series</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td></td>
<td>BIZ-UDMinchoM.ttc</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>BIZ-UDGothicR.ttc</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>BIZ-UDGothicB.ttc</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>BIZ-UDGothicB.ttc</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></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</th>
<th>series</th>
<th>morisawa-pro</th>
<th>morisawa-pr6n</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>medium</td>
<td>A-OTF-RyuminPro-Light.otf</td>
<td>A-OTF-RyuminPr6N-Light.otf</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>A-OTF-FutoMinA101Pro-Bold.otf</td>
<td>A-OTF-FutoMinA101Pr6N-Bold.otf</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>A-OTF-GothicBBBPro-Medium.otf</td>
<td>A-OTF-GothicBBBPr6N-Medium.otf</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>A-OTF-FutoGoB101Pro-Bold.otf</td>
<td>A-OTF-FutoGoB101Pr6N-Bold.otf</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>A-OTF-MidashiGoPro-MB31.otf</td>
<td>A-OTF-MidashiGoPr6N-MB31.otf</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>A-OTF-Jun101Pro-Light.otf</td>
<td>A-OTF-ShinMGoPr6N-Light.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</th>
<th>series</th>
<th>yu-win</th>
<th>yu-win10</th>
<th>yu-osx</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>light</td>
<td>YuMincho-Light</td>
<td>YuMincho-Light</td>
<td>(YuMincho Medium)</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>YuMincho-Regular</td>
<td>YuMincho-Regular</td>
<td>YuMincho Medium</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>YuMincho-DemiBold</td>
<td>YuMincho-DemiBold</td>
<td>YuMincho DemiBold</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>YuGothic-Regular*</td>
<td>YuGothic-Regular*</td>
<td>YuGothic Medium*</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic Bold</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic-Bold</td>
<td>YuGothic Bold</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></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</th>
<th>series</th>
<th>default, 90jis option</th>
<th>jis2004 option</th>
</tr>
</thead>
<tbody>
<tr>
<td>mincho</td>
<td>medium</td>
<td>Moga90Mincho</td>
<td>MogaMincho</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Moga90Mincho Bold</td>
<td>MogaMincho Bold</td>
</tr>
<tr>
<td>gothic</td>
<td>medium</td>
<td>Moga90Gothic</td>
<td>MogaGothic</td>
</tr>
<tr>
<td></td>
<td>bold</td>
<td>Moga90Gothic Bold</td>
<td>MogaGothic Bold</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>Moga90Gothic Bold</td>
<td>MogaGothic Bold</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></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.

**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>bold</td>
<td>Source Han Sans Medium</td>
<td>Source Han Sans JP Medium</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>Source Han Sans Heavy</td>
<td>Source Han Sans JP Bold</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>Source Han Sans Heavy</td>
<td>Source Han Sans JP Heavy</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>bold</td>
<td>Noto Sans CJK Medium</td>
<td>Noto Sans CJK JP Medium</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>Noto Sans CJK Bold</td>
<td>Noto Sans CJK JP Bold</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>Noto Sans CJK Black</td>
<td>Noto Sans CJK JP Black</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>bold</td>
<td>HaranoAjiGothic-Bold.otf</td>
</tr>
<tr>
<td></td>
<td>extra bold</td>
<td>HaranoAjiGothic-Heavy.otf</td>
</tr>
<tr>
<td>rounded gothic</td>
<td></td>
<td>HaranoAjiGothic-Heavy.otf</td>
</tr>
</tbody>
</table>

### 11.6.3 Presets which do not support multi weights

Next, we describe settings for using only single weight.

<table>
<thead>
<tr>
<th>noembed</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>
</tr>
<tr>
<td>gothic</td>
<td>GothicBBB-Medium (non-embedded)</td>
<td>IPA Gothic</td>
<td>IPAex Gothic</td>
</tr>
</tbody>
</table>
11.6.4 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>ns-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></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>
<tr>
<td>rounded gothic</td>
<td></td>
<td>HG Maru Gothic M PRO</td>
<td>HG Maru Gothic M PRO</td>
<td>HG Maru Gothic M PRO</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.).

jis0, 90jis by file name (hgrme.ttc, hgrge.ttc, hgrsgu.ttc, hgrsmp.ttf).


11.6.5 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\{⟨name⟩\}\{(⟨specification⟩)\}
Define new preset ⟨name⟩. This ⟨name⟩ cannot be same as other presets, options described in Subsubsection 11.6.1, nor following 11 strings:

mc mc-l mc-m mc-b mc-bx gt gt-m gt-b gt-bx gt-eb mg-m

⟨specification⟩ is a comma-separated list which consists of other presets and/or the following keys:

mc-⟨font⟩ mincho light
mc-⟨font⟩ mincho medium
mc-⟨font⟩ mincho bold
mc-⟨font⟩ synonym for mc-b=⟨font⟩
gt-⟨font⟩ gothic medium
gt-⟨font⟩ gothic bold
gt-⟨font⟩ synonym for gt-b=⟨font⟩
gt-eb=⟨font⟩ gothic extra bold
mg-⟨font⟩ rounded gothic
mc=⟨font⟩ Same as
mc-l=⟨font⟩, mc-m=⟨font⟩, mc-b=⟨font⟩
gt=⟨font⟩ Same as
gt-m=⟨font⟩, gt-b=⟨font⟩, gt-eb=⟨font⟩

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

\ltjnewpreset*\{⟨name⟩\}\{(⟨specification⟩)\}
Almost same as \ltjnewpreset. However, if ⟨name⟩ 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:

\begin{verbatim}
\texttt{\textbackslash ltjnewpreset\{hoge\}\{piyo,mc-b=HiraMinProN-W6\}}
\texttt{\textbackslash ltjnewpreset\{piyo\}\{mg-m=HiraMaruProN-W4\}}
\texttt{\textbackslash ltjapplypreset\{hoge\}}
\end{verbatim}

\begin{itemize}
\item Restrictions
\end{itemize}

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 90jis or jis2004 is specified. This feature is not usable in presets which are defined by \texttt{\textbackslash ltjnewpreset}.
\end{itemize}
Part III
Implementations

12 Storing Parameters

12.1 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 \Q = 0.25 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 \dimen register as \jQ.

\ltj@dimen@zu (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@curftnt (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} pt).

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

\ltj@tablshift (attribute)
\ltj@tkblshift (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:

italic (1) Kerns from italic correction (/\), or from kerning information of a Japanese font. These kerns are “ignored” in the insertion process of JAgue, unlike explicit \kern.

packed (2)

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

from_jfm (from_jfm + 63) (4–67) Glues/kerns from JFM.
kanji_skip (68), kanji_skip_jfm (69) Glues from kanjiskip.
xkanji_skip (70), xkanji_skip_jfm (71) Glues from xkanjiskip.
processed (73) Nodes which is already processed by ....
ic processed (74) Glues from an italic correction, but already processed in the insertion process of \texttt{JAglues}.

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

\texttt{\textbackslash ltj@kcat i} (attribute) Where \textit{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 \texttt{JAchars}.

\texttt{\textbackslash ltj@dir} (attribute) \texttt{dir_node_auto} (128)
\texttt{dir_node_manual} (256)
\texttt{\textbackslash ltjlineendcomment} (counter)

Furthermore, \texttt{LuATeX-ja} uses several user-defined whatsit nodes for internal processing. All those nodes except \texttt{direction} whatsits store a natural number (hence its type is 100). \texttt{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}.

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

\texttt{stack_marker} Nodes for \texttt{LuATeX-ja}’s stack system (see the next subsection). The value field of these nodes is current group level.

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

\texttt{replace_vs} Similar to \texttt{char_by_cid} whatsits above. These nodes are for \texttt{ALchar} which the callback process of \texttt{luaotfload} won’t be applied.

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

\texttt{direction}

These whatsits will be removed during the process of inserting \texttt{JAglues}.

### 12.2 Stack system of \texttt{LuATeX-ja}

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

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

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

```latex
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;
  d = box\_max\_depth;
  unsave();
}
```

52
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\textsuperscript{14}.

These are two \TeX count registers for maintaining information: \\texttt{\l@t@stack} for the stack level, and \\texttt{\l@t@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 \texttt{\l@tjsetparameter}, 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 whatsis 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 \\texttt{\l@tj@stack} and \\texttt{\l@tj@group@level} have to be local always, regardless the value of \texttt{\globaldefs}. To solve this problem, we use another trick: the assignment \texttt{\directlua{tex.globaldefs=0}} is always local.

12.3 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:
\texttt{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 \texttt{luatexja.isglobal} (stored globally if and only if \texttt{luatexja.isglobal == 'global')}.

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

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

12.4 Extending Parameters

Keys for \texttt{\l@tjsetparameter} and \texttt{\l@tjgetparameter} can be extended, as in luatexja-adjust.\textsuperscript{14}[Dev-luatex] tex.currentgrouplevel, a post at 2008/8/19 by Jonathan Sauer.
\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\}%

\begin{figure}[htb]
\centering
\includegraphics[width=\textwidth]{figure10}
\caption{Definition of parameter setting commands}
\end{figure}

\subsection*{Setting parameters}
\ref{fig:setting-parameters} 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 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{\texttt{luatexja.isglobal}} = \begin{cases} 
'global' & \text{(global assignment),} \\
'local' & \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}.

\subsection*{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 \texttt{string} 42.
\begin{verbatim}
function luatexja.unary\_pars.hoge (t)
  return 42
end
\end{verbatim}

Here the argument of \texttt{luatexja.unary\_pars.hoge} is the stack level of Lua\TeX\-ja's stack system (see Subsection 12.2).

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}
function luatexja.binary\_pars.fuga (c, t)
  return tostring(c) .. ', ' .. tostring(42)
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}
\ltj@@decl@array@param{fuga}
\end{verbatim}
to indicate that "the parameter fuga needs an additional argument".

\section{Linebreak after a Japanese Character}
\subsection{Reference: behavior in \texttt{p\TeX}}
In \texttt{p\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
p\TeX's input processor can be described in terms of a finite state automaton, as that of \TeX in Section 2.5 of [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 11. Note that p\TeX doesn’t leave state $K$ after “beginning/ending of a group” characters.

### 13.2 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 l\textbackslash tj1\lineendcomment} is appended to an input line, before Lua\TeX actually process it, if and only if the following three conditions are satisfied:

1. The category code of \texttt{\endlinechar} is 5 (end-of-line).
2. The category code of \texttt{\backslash l\textbackslash tj1\lineendcomment} itself is 14 (comment).

\texttt{\backslash l\textbackslash tj1\lineendcomment} usually, it is (return) (whose character code is 13).
3. The input line matches the following "regular expression":

\((\text{any char})^* (\text{JAchar})(\{\text{catcode} = 1\} \cup \{\text{catcode} = 2\})^*\)

\begin{itemize}
  \item \textbf{Remark} \ The following example shows the major difference from the behavior of \TeX.
\end{itemize}

\begin{verbatim}
\fontspec[Ligatures=TeX]{Linux Libertine O}
\ltjsetparameter{autoxspacing=false}
\ltjsetparameter{jacharrange={-6}}xあ
y\ltjsetparameter{jacharrange={+6}}zい
\end{verbatim}

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 \textbf{ALchar}.

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

\begin{itemize}
  \item When line 3 is processed by \texttt{process_input_buffer} callback, “あ” is considered as an \textbf{JAchar}. Since line 3 ends with an \textbf{JAchar}, the comment character (whose character code is \texttt{\ltjlineendcomment}) is appended to this line, and hence the linebreak immediately after this line is ignored.
  \item When line 4 is processed by \texttt{process_input_buffer} callback, “い” is considered as an \textbf{ALchar}. Since line 4 ends with an \textbf{ALchar}, the linebreak immediately after this line emits a space.
\end{itemize}

14 \textbf{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 \TeX series, there is no method to make Japanese characters active; a patch \texttt{jlisting.sty} ([4]) resolves the problem forcibly.

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

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

14.1 \textbf{Notes and additional keys}

\begin{itemize}
  \item \textbf{Variation selectors} \ lltjp-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 lltjp-listings is loaded at \texttt{\begin{document}}.
    \begin{itemize}
      \item \texttt{vsraw} is a key which takes a boolean value, and its default value is false.
        \begin{itemize}
          \item If the \texttt{vsraw} key is true, then variation selectors are "combined" with the previous character.
          \begin{verbatim}
          \begin{lstlisting}[vsraw=true]
          城市, 葛飾区, 葛西
          \end{lstlisting}
          \end{verbatim}
        \end{itemize}
      \item 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}[vsraw=false, vscmd=\ltjlistingsvsstdcmd]
        城市, 葛飾区, 葛西
        \end{lstlisting}
        \end{verbatim}
        \begin{verbatim}
        城市, 葛飾区, 葛西
        \end{verbatim}
        \end{itemize}
    \end{itemize}
\end{itemize}
For example, the following code is the setting of the vscmd key in this document.

```latex
\def\IVSA#1#2#3#4#5{% 
\hbox to1em{\hss\textcolor{blue}{\raisebox{3.5pt}{\normalfont\ttfamily \fboxsep=0.5pt\fbox{\hbox to0.75em{\hss\tiny \oalign{0#1#2\crcr#3#4#5\crcr}\hss}}}}}\hss}
\% 
\}
\{\catcode`%=11
\gdef\IVSB#1{\expandafter\IVSA\directlua{
local cat_str = luatexbase.catcodetables['string']
tex.sprint(cat_str, string.format('%X', 0xE00EF+#1))}}
}\lstset{vscmd=\IVSB}
```

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

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

```
: H :
: H H H H :
```

lltjp-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.

```
: H :
: H H H H :
```

### 14.2 Class of characters

Roughly speaking, the listings package processes input as follows:

1. Collects letters and digits, which can be used for the name of identifiers.
2. When reading an other, outputs the collected character string (with modification, if needed).
3. Collects others.
4. When reading a letter or a 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 lltjp-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>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>\lst@ifletter</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>\lst@ifkanji</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

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

| Meaning | |
|---------| |
| most of Japanese char | opening brackets | closing brackets |

57
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></td>
<td></td>
<td>collects</td>
<td>outputs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>outputs</td>
<td>collects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kanji</td>
<td></td>
<td>outputs</td>
<td></td>
<td>collects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>collects</td>
<td></td>
</tr>
<tr>
<td>Close</td>
<td></td>
<td>outputs</td>
<td></td>
<td></td>
<td></td>
<td>collects</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 U+0080 *except Variation Selectors* are classified into above 5 classes by the following rules:

- **ALchars** above or equal to U+0080 are classified as Letter.
- **JAchars** 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 (U+FF61–U+FF9F) is same as the width of **ALchar**; the width of the other **JAchars** is double the width of **ALchar**.

This classification process is executed every time a character appears in the 1st `lstlisting` environment or other environments/commands.

### 15 Cache Management of LuaT\TeX-ja

LuaT\TeX-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) `$TEXMFVAR/luatexja`.
- In addition to caches of the text form (the extension is “.lua”), caches of the *binary*, precompiled form are supported.
  - In loading a cache, the binary cache precedes the text form.
  - When LuaT\TeX-ja updates a cache `hoge.lua`, its binary version is also updated.

#### 15.1 Use of cache

LuaT\TeX-ja uses the following cache:

`ltj-cid-auto-adobe-japan1.lua`

The font table of a CID-keyed non-embedded Japanese font. This is loaded in every run. It is created from three CMaps, `UnI1S2004-UTF32-{H,V}` and `Adobe-Japan1-UCS2`, and this is why these two CMaps are needed in the first run of LuaT\TeX-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>
<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>ltj-cid-auto-adobe-japan1.lua</td>
<td>UniJIS2004-UTF32-*</td>
</tr>
<tr>
<td>Adobe-Korea1-*</td>
<td>ltj-cid-auto-adobe-korea1.lua</td>
<td>UniKS-UTF32-*</td>
</tr>
<tr>
<td>Adobe-GB1-*</td>
<td>ltj-cid-auto-adobe-gbl.lua</td>
<td>UniGB-UTF32-*</td>
</tr>
<tr>
<td>Adobe-CNS1-*</td>
<td>ltj-cid-auto-adobe-cns1.lua</td>
<td>UniCNS-UTF32-*</td>
</tr>
</tbody>
</table>

`ltj-jisx0208.{luc|lub}`

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 \( p\TeX \).

### 15.2 Internal

Cache management system of Lua\TeX-ja is stored in `luatexja.base` (`ltj-base.lua`). There are three public functions for cache management in `luatexja.base`, where \( \langle \text{filename} \rangle \) stands for the file name without suffix:

- `save_cache(\langle \text{filename} \rangle, \langle \text{data} \rangle)`
  
  Save a non-nil table \( \langle \text{data} \rangle \) into a cache \( \langle \text{filename} \rangle \). Both the text form \( \langle \text{filename} \rangle .lua \) and its binary version are created or updated.

- `save_cache_luc(\langle \text{filename} \rangle, \langle \text{data} \rangle[, \langle \text{serialized_data} \rangle])`
  
  Same as `save_cache`, except that only the binary cache is updated. The third argument \( \langle \text{serialized_data} \rangle \) is not usually given. But if this is given, it is treated as a string representation of \( \langle \text{data} \rangle \).

- `load_cache(\langle \text{filename} \rangle, \langle \text{outdate} \rangle)`
  
  Load the cache \( \langle \text{filename} \rangle \). \( \langle \text{outdate} \rangle \) 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 \( \langle \text{filename} \rangle .\{\text{luc|lub}\} \). 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 text form \( \langle \text{filename} \rangle .lua \). Hence, the return value of `load_cache` is non-nil, if and only if the updated cache is found.
References