

# ルビ実装のサンプル

2014年2月8日

## 1 使い方

- パッケージ読み込み：`\usepackage{luatexja-ruby}`のみ。L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>のみサポート。
- 命令：
  - `\ltjruby[option]{親|文|字}{おや|も|じ}`
  - `\ltjsetruby{option}`
- `\ltjsetParameter` に追加のキー
  - `rubypreintrusion={char_code, pre_int}`：全角単位
  - `rubypostintrusion={char_code, post_int}`：全角単位
- `\ltjsetruby{option}`他でサポートするオプション：後ろに指定したものが優先
  - `mode`：進入処理のモードを表す bit vector。意味は `luatexja-ruby.sty` 参照。
  - `intrusionpre=<real>`：前進入量指定。ルビ全角単位で指定。負の長さは文字種に応じた自動指定を意味。
  - `intrusionpost=<real>`：後進入量指定。
  - `stretchhead={left}{middle}{right}`：行頭形でルビ文字の合計長が親文字の合計長より長い時に、親文字間に入れる空白の割合。それぞれ 0-7 の自然数で指定。
  - `stretchend={left}{middle}{right}`：行末形～。
  - `stretch={left}{middle}{right}`：行中形～。
  - `stretchruby={left}{middle}{right}`：ルビ文字の方が長い時の空白の割合。
  - `maxmargin=<real>`：親文字の方がルビより長い時に、ルビの先頭と親文字の先頭に許される最大の空白量。親文字全角単位。
  - `rubysize=<real>`：ルビ文字の親文字に対する大きさ。

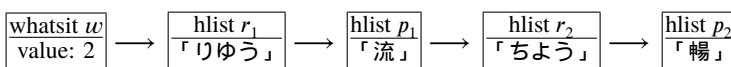
## 2 実装の大まかな方法

次の例で説明する。  
りゆうちよう

.....を流暢に.....

.....を\ruby{流|暢}{りゆう|ちよう}に.....

1. \ruby コマンド自体は、一旦次の node list を値とする whatsit  $W$  を作って、現在の水平リストへと挿入する（必要ならば\leavevmode も実行）:



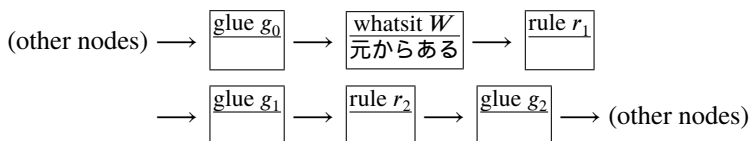
ここで、最初の  $w$  の値 2 は、ルビが 2 つのパーツ「流」「暢」からなっていることを表している。 $r_i$  達の中の文字は既にルビの大きさである。この値を  $cmp$  とおこう。

2. LuaTeX-ja の和文処理グルー挿入処理においては、「先頭が『流』、最後が『暢』であるような hbox を\unhbox で展開したもの」と扱われる。言い換えれば、ルビ部分を無視した単なる「流暢」という和文文字の並びとして扱われる\*1。次のサンプルを参照

.A  
.A

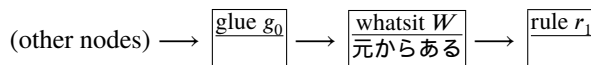
```
1 \leavevmode\hbox{ JA\
2 % xkanjiskip
3 \ruby{ J}{A}
4 % 2分
```

3. 和文処理グルーの挿入が終わった後で、可動グルーのルビのためのノードの挿入に入る。  
(a)  $W$  の周辺は次のようなノード列になる。



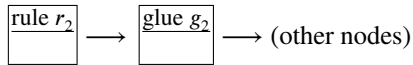
$W$  の前後に  $2cmp + 1 = 5$  個のノードが挿入される。

- (b) このようにノードを挿入する目的は、TeX の行分割処理自体に影響を加えずに可動グルーのルビを実現させることにある。



行の境目

\*1 「流」「暢」の間のグルーは既に入っている、と扱われる。



のようになったとしたら、「流」「暢」の間で行分割が起きた、ということがわかり、 $g_i, r_i$  達のノードを適切に置き換えればよい（後で詳しく説明する）。

(c) なお、 $r_i$  達の高さ・深さは組み上がった後のそれであり、現状では親文字部分の高さ・深さと一致する。 $g_i, r_i$  達の幅は、以下の対応に沿って計算される。

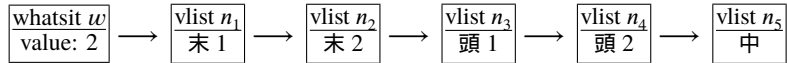
node 名	組み方	サンプル	対応するノード並び
$n_1$	行末 1 グループ	$\boxed{\text{を}} \overset{\text{りゆう}}{\text{流}}$	$g_0 \rightarrow W \rightarrow r_1$
$n_2$	行末 2 グループ	$\boxed{\text{を}} \overset{\text{りゆう}}{\text{流}} \overset{\text{ちよう}}{\text{暢}}$	$g_0 \rightarrow W \rightarrow r_1 \rightarrow g_2 \rightarrow r_2$
$n_3$	行頭 1 グループ	$\overset{\text{ちよう}}{\text{暢}} \boxed{\text{に}}$	$r_2 \rightarrow g_2$
$n_4$	行頭 2 グループ	$\overset{\text{りゆう}}{\text{流}} \overset{\text{りゆうちよう}}{\text{暢}} \boxed{\text{に}}$	$W \rightarrow r_1 \rightarrow g_2 \rightarrow r_2 \rightarrow g_2$
$n_5$	行中	$\boxed{\text{を}} \overset{\text{りゆう}}{\text{流}} \overset{\text{りゆうちよう}}{\text{暢}} \boxed{\text{に}}$	$g_0 \rightarrow W \rightarrow r_1 \rightarrow g_2 \rightarrow r_2 \rightarrow g_2$

この 5 元連立方程式を解いて  $g_i, r_i$  達の幅を求める。なお、進入長は予め計算しておく。例えばこの場合、

$$g_0 + r_1 + g_2 + r_2 + g_2 = 3zw - (0.25zw \times 2) = 2.5zw$$

である ( $zw$  は親文字全角の幅)。

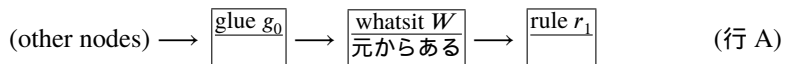
(d) また、ルビ処理を統括している whatsit  $W$  の値も



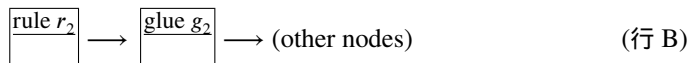
に置き換えておく。

4. LuaTeX の行分割処理を普通に行う。
5. 行分割の結果に従って、 $g_i, r_i$  達を適切に置換する。

例えば行分割の結果

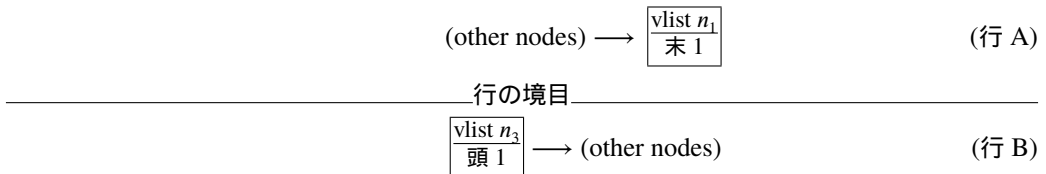


行の境目



のようになったとしよう。

- (a) 処理は段落の上の行から順番に行われる．行 A の処理がまわってきたとしよう．
- (b) 行 A の先頭から順番に眺めていく．すると「whatsit  $W$  由来」のノード， $g_0, W, r_1$  が見つかり，行 A はここで終わっている．  
 まず，行 A の hbox の中身から whatsit  $W$  を消去（リストから取り除くだけで， $W$  のメモリを解放するわけではない）する． $g_0, (W), r_1$  というノードの並びは，「行末 1 グループ」 $n_1$  に対応しているので， $g_0, r_1$  を行 A から除去・メモリ解放し，代わりに  $n_1$  を行 A の中身に追加する．
- (c) 次に行 B の処理にうつる．行 A でルビの処理は完了していない（2 パーツのルビなのにまだ 1 パーツ目しか使っていないからである）ので，「whatsit  $W$  由来」のノードがいくつか残っているはずである．  
 案の定， $r_2, g_2$  というノード列が見つかった．これは「行頭 1 グループ」 $n_3$  に対応しているので， $r_2, g_2$  を行 B から除去・メモリ解放し，代わりに  $n_3$  を行 B の中身に挿入する．
- (d) これで 2 パーツとも使い切ったことになるので，隔離しておいた  $W$  を，（使われなかった  $n_2, n_4, n_5$  などと共に）メモリ解放する．結果として次のようになった：



### 3 TODO

- 熟語ルビの配置が「モノルビの連続」「グループルビ」のどちらかしかない（行分割は可能）。「一字だけかけても良い」なんてどうやるんだろう．
- 図 3.82 のように，ルビ文字のみ出しが繋がらないようにする処理が不安定である．例えば，

りょう    みささぎ  
 陵と 陵

において，後者の「<sup>みささぎ</sup>陵」のルビは前の「と」にかかっていない．これは次の理由によるものである：

1. ルビのところで行分割を正しく実装するためには，行分割前に行頭・行中・行末形の幅を決定しておかねばならない．
2. 「<sup>りょう</sup>陵」の後側進入量は，行頭形するとき（ルビ全角）が最大となる．

3. 「<sup>みさきぎ</sup>陵」のルビ文字のはみ出しが重ならないように、前側進入許容量は

$$\underbrace{0.5 \text{ zw}}_{\text{元々の許容量}} - \underbrace{0.5 \text{ zw}}_{\text{前のルビの後側進入量の最大値}} = 0$$

となる。そのため、「<sup>みさきぎ</sup>陵」のルビ文字は「と」にはかからない。

1 パスでは厳しいだろうが、2 パスでなら全てのルビにおいて「どの形で出力されたか」を aux に保存することで改善されると思う。しかし、どのように保存するか（具体的には、各 ruby 命令の呼びだしをどういうキーに変換するか）が問題だ。

- 実装方法の都合上、ルビの直前・直後・途中で 2 箇所以上の改行が起きる場合（以下のパーツの組み方が出てくる）に対応できない。

組み方	サンプル
単独 1	りゆう 流
単独 2	ちよう 暢
単独 (1 + 2)	りゆうちよう 流 暢

進入 (intrusion) と突出 (protrusion)

進入ありの例： <sup>あかつき</sup>と <sup>あかつき</sup>暁の

進入なしの例： と <sup>あかつき</sup>暁 の

突出ありの例： <sup>ちようしゆう</sup>聴 <sup>ちようしゆう</sup>衆

突出なしの例： <sup>ちようしゆう</sup>聴 衆

mode

mode 00xx: <sup>うつく</sup>は美しい <sup>ちようしゆう</sup>は聴衆と

mode 01xx: <sup>うつく</sup>は美しい <sup>ちようしゆう</sup>は聴衆と

mode 10xx: <sup>うつく</sup>は美しい <sup>ちようしゆう</sup>は聴衆と

mode 11xx: <sup>あかつき</sup>は暁立 <sup>ちようしゆう</sup>は聴衆立

ここでは「立」には 0.5 字がけを許容  
中付きでは 00xx, 11xx を選択して下さい

#### 4 いくつかの例

ゴールデンゲートブリッジ  
 ああああ黄金橋こいうえおかきくけこあ  
ゴールデンゲートブリッジ 黄金橋こいうえおかきくけこあ ゴールデンゲートブリッジ 黄金橋  
ゴールデンゲートブリッジ いうえおかきくけこあ ゴールデンゲートブリッジ 黄金橋こいうえおか  
ゴールデンゲートブリッジ きくけこあ ゴールデンゲートブリッジ 黄金橋こいうえおかきくけこあ  
ゴールデンゲートブリッジ 黄金橋こいうえおかきくけこ

こあ国府津こいうえおかきくけこあ国府津  
こいうえおかきくけこあ国府津こいうえおか  
こきくけこあ国府津こいうえおかきくけこあ  
こ国府津こいうえおかきくけこあ国府津こいう  
こえおかきくけこ

あ いうえおかきくけこあ  
 いうえおかきくけこあ いうえお  
 かきくけこあ いうえおかきくけこ  
 イあ いうえおかきくけこあ  
 いうえおかきくけこウあ いう  
 えおかきくけこエあ いうえおか  
 きくけこあ いうえおかきくけこあ  
 いうえおかきくけこ

あ いうえおかきくけこあ い  
 うえおかきくけこあ いうえおかきく  
 けこあ いうえおかきくけこイあ  
 いうえおかきくけこあ いうえお  
 かきくけこウあ いうえおかきくけこ  
 エあ いうえおかきくけこあ い  
 うえおかきくけこあ いうえおかきく  
 けこ

いよう あ異様いよういうえくけこあ異様いよういうえくけこ  
いよう あ異様いよういうえくけこあ異様いよういうえくけこイ  
いよう あ異様いよういうえおかきくけこ

あ いうえおかきくけこうえおかきくけこ  
あ いうえおかきくけこうえおかきくけこあ  
いう おかきくけこうえおかきくけこあ  
いう おかきくけこうえおかきくけこあ  
いう おかきくけこうえおかきくけこあ  
いうえおかきくけこ

標準 又 <sup>うけたまわ</sup>承<sup>そ</sup>る 疎 <sup>にわか</sup>は俄 勉強 <sup>にわか</sup>後俄 勉強  
肩つき 又 <sup>うけたまわ</sup>承<sup>そ</sup>る 疎 <sup>にわか</sup>は俄 勉強 <sup>にわか</sup>後俄 勉強  
拡張肩つき 又 <sup>うけたまわ</sup>承<sup>そ</sup>る 疎 <sup>にわか</sup>は俄 勉強 <sup>にわか</sup>後俄 勉強

## 5 jlreq 20120403 の例

### 3.3.1 節

3.49 君子<sup>くんし</sup>は和<sup>わ</sup>して同<sup>どう</sup>ぜず

3.50 人<sup>ひと</sup>に誨<sup>おし</sup>えて倦<sup>う</sup>まず

3.51 鬼門<sup>きもん</sup>の方角<sup>ほうかく</sup>を凝視<sup>ぎょうし</sup>する

3.52 鬼門<sup>きもん</sup>の方角<sup>ほうかく</sup>を凝視<sup>ぎょうし</sup>する

3.53 茅場町<sup>かやばちよう</sup> 茅場町<sup>かやばちよう</sup>

3.54 紫陽花<sup>あじさい</sup> 垣塙<sup>るつほ</sup> 田舎<sup>いなか</sup>

3.55 模型<sup>モデル</sup> 顧客<sup>クライアント</sup> 境界面<sup>インターフェース</sup> 避難所<sup>アジール</sup>

3.56 編集者<sup>editor</sup> エディター<sup>editor</sup>

### 3.3.3 節

3.58 に幟<sup>のぼり</sup>を に幟<sup>のぼり</sup>を 章編三絶<sup>いへんさんぜつ</sup>

3.59 に幟<sup>のぼり</sup>を

3.60 章編三絶<sup>いへんさんぜつ</sup> 章編三絶<sup>いへんさんぜつ</sup>

### 3.3.4 節

3.61 図 ( 両側ルビ ) はまだ未サポートにより省略

### 3.3.5 節 モノルビ

3.62 の薬<sup>やく</sup>に

3.63 版面<sup>ち</sup>の地<sup>ち</sup>に 版面<sup>ち</sup>の地<sup>ち</sup>に

3.64 × 版面<sup>ち</sup>の地<sup>ち</sup>に ( まだ縦組み未サポート )

3.65 の砦<sup>とりで</sup>に の砦<sup>とりで</sup>に

3.66 上<sup>しゆん</sup>の旬<sup>しゆん</sup>に 後旬<sup>しゆん</sup>に

3.66 下<sup>しゆん</sup>の旬<sup>しゆん</sup>又 後旬<sup>しゆん</sup>又

### 3.3.6 節 グループルビ

3.67 は冊子<sup>コーデックス</sup>体と

3.68 模型<sup>モデル</sup> 利用許諾<sup>ライセンス</sup>

3.69 模型<sup>モデル</sup> 利用許諾<sup>ライセンス</sup>

3.70 なげきの聖母像<sup>ビエタ</sup> 自動調整

3.71 顧客<sup>クライアント</sup> 境界面<sup>インターフェース</sup>

3.72 顧客<sup>クライアント</sup> 境界面<sup>インターフェース</sup>



3.3.7 節 熟語ルビ

3.73 左 杞憂 畏怖

3.73 右 杞憂 畏怖

3.74 の流儀を の無常を の成就を

3.74 の紋章を の象徴を

3.75 の流儀を の無常を の成就を

3.75 の紋章を の象徴を

3.76 x の流儀を の無常を

要調整 3.77	流儀を	..... の	.....の無
			常を

3.3.8 節 ルビはみ出し

3.78 人は死して名を残す

要調整 3.79 漢字の部首には偏・冠・脚・旁がある

要調整 3.79 漢字の部首には偏，冠，脚，旁がある

3.79 この噂の好きな人は懐ぐあいもよく、檜を

3.80 漢字の部首には「偏」「冠」「脚」「旁」がある

3.80 この噂好きな人は懐 具合もよく、檜 材を

要調整 3.81 x に暁の趣を

3.82 に暁の 趣を

3.83 この噂の好きな人は懐ぐあいもよく、檜を

3.83 この噂好きな人は懐 具合もよく、檜材を

3.84 この噂の好きな人は懐 ぐあいもよく、檜 を

3.84 この噂好きな人は 懐 具合もよく、檜 材を

要調整 3.85	あああああああああ	あいの	径
	あああああああああ	あいの	径
	を	あああああああああ	ああ

3.86	径を	あああああああああ	ああ
		あああああああああ	ああ

要調整 3.87	あああああああああ	ああ共	飾り
	あああああああああ	ああ共	飾り

## 要調整

### F.1--2 節

F.01 治癒ち ゆ 模索も さく 遷移せん い 混沌こん どん

F.01 中治癒ち ゆ 模索も さく 遷移せん い 混沌こん どん

F.02 橋頭堡きょうとうぼ

F.03 凝視ぎょうし 調理師ちようりし 思春期ししゆんき 管状花かんじょうか 蒸気船じょうきせん

F.03 凝視ぎょうし 調理師ちようりし 思春期ししゆんき 管状花かんじょうか 蒸気船じょうきせん

F.04 未熟みじゆく 法華経ほけきよう 顕微鏡けんびきよう 課徴金かちようきん 古戦場こせんじよう

F.04 未熟みじゆく 法華経ほけきよう 顕微鏡けんびきよう 課徴金かちようきん 古戦場こせんじよう

F.05 の峻別しゆんべつは

F.06 以降は未チェック