

pxbabel パッケージ

八登崇之 (Takayuki YATO; aka. “ZR”)

v1.4 [2023/02/25]

1 概要

(u)pL^AT_EX における日本語・簡体字中国語・繁体字中国語・韓国語の間の「和文 (CJK) フォント」の切り替えを Babel のインタフェースを用いて行えるようにする。

■前提環境

- T_EX フォーマット：pL^AT_EX / upL^AT_EX (拡張エンジンを含む)

■依存パッケージ

- babel パッケージ
- japanese-otf パッケージ (pL^AT_EX の場合)

1.1 用語集

- CJK 言語：「日本語」「簡体字中国語」「繁体字中国語」「韓国語」の総称。^{*1}
- CJK 言語識別子：本パッケージにおいて CJK 言語を指し示す、以下の文字列。
 - `japanese`：日本語
 - `korean`：韓国語
 - `schinese`：簡体字中国語
 - `tchinese`：繁体字中国語
- 言語名：Babel で指定する (`\foreignlanguage` 等の引数にする) 言語を表す文字列。
※既定では CJK 言語に対する「言語名」は「CJK 言語識別子」と同じ (例えば日本語の「言語名」は `japanese`) であるが、そうでない設定も可能である。
- 仮想言語：本パッケージの動作により自動的に定義される、通常の言語オプション定義 (キャプション文字列) を伴わない Babel 言語のこと。^{*2}本パッケージでは CJK 言語を仮想言語として扱えるため、各言語の定義ファイルを用意する必要がない。

^{*1} 普通の意味で「言語」の区別といえるかは問題にしない。

^{*2} つまり、言語オプション定義ファイル (`*.ldf`) の読込により定義される Babel 言語は仮想言語ではない。最近の Babel では「`ini` ファイルにより言語を定義する」という機能があるが、そのように定義された Babel 言語も仮想言語ではない。

※ CJK 言語を仮想言語にせず定義ファイルを利用することも可能。

2 読込

0.9 版より babel パッケージと一緒に読み込む方式（一括読込モード）をサポートした。従来通り、別個に読み込む方式（分離読込モード）も可能である。どちらであるかは、pxbabel 読込時に babel が読込済かどうかで判別する。

〈一括読込モードの場合〉

```
\usepackage[〈オプション〉]{pxbabel}
```

〈分離読込モードの場合〉

```
\usepackage[〈Babel のオプション〉]{babel}
```

```
\usepackage[〈オプション〉]{pxbabel}
```

2.1 パッケージオプション

- 〈CJK 言語識別子〉=(言語名)： CJK 言語を仮想言語として扱い、実際に使用する Babel の言語名を指定する。
※詳細は 3 節を参照。
※既定で用いられる言語名は CJK 言語識別子と同一の文字列である。
- 〈CJK 言語識別子〉： 「〈CJK 言語識別子〉=(CJK 言語識別子)」と同じ。
※既定でこれと同等の設定が適用されるので通常は指定する必要がないが、一括読込モードで「最後の言語（基底言語になる）」を指定するのに有用である。
- 〈CJK 言語識別子〉=+(言語名)： CJK 言語を非仮想言語として扱い、実際に使用する Babel の言語名を指定する。
※「〈言語名〉.ldf」という名前の言語オプション定義ファイルが読み込まれる。
- +(CJK 言語識別子)： 「〈CJK 言語識別子〉=+(CJK 言語識別子)」と同じ。
- Babel の有効な言語名： その言語の定義ファイルを読み込む。
※分離読込モードでの非 CJK 言語の読込については確実に動作させるのが困難であるため、1.3 版において**非推奨**の扱いとする。すなわち、分離読込モードでは CJK 言語のみが指定可能となり、CJK 言語以外は babel のオプションとして指定する必要がある。
- main=〈言語名〉： 文書開始時に有効となる基底言語を、言語名*³で指定する。
※ main の指定がない場合、分離読込モードでは babel の読込時に基底言語であったものがそのまま基底言語になる。一括読込モードでは pxbabel で最後に読み込まれた言語が基底言語になる。
- caption=〈値〉： キャプション言語（キャプションや日付の文字列で使われる言語）の固定に関する設定を行う。有効な値は次の通り：
 - 有効な言語名： その言語に固定する。
 - main： 文書開始時の基底言語に固定する。

*3 CJK 言語について言語名と CJK 言語識別子が異なる設定を用いる場合、main に指定するのは言語名の方であることに注意。

- `default`： キャプション文字列を文書クラスで指定されたものから変更しない。
 - `switch`： キャプション言語の固定を行わない。(Babel の既定動作。)
 - `auto` (既定値)： 文書開始時の基底言語が仮想言語の場合は `default`、それ以外は `switch`。
- ※詳細は 4 節を参照。
- `noswitchcaption`： `caption=default` と同じ。
 - `switchcaption`： `caption=switch` と同じ。
 - ※ `(no)switchcaption` は 0.9 版より前との互換性のため存在する。
 - `switchfont` (既定値)： CJK 言語でフォントを切り替える機能を有効にする。
 - `noswitchfont`： `switchfont` の否定。
 - `stdjaencoding`： 入力言語が日本語の場合には `upLATEX` 標準の「和文エンコーディング」を適用する。
 - ※詳細は 5.1 節を参照。 `upLATEX` でのみ意味をもつ。
 - `nostdjaencoding` (既定)： `nostdjaencoding` の否定。
 - ※ `japanese-otf` パッケージが読込済の場合には既定が `stdjaencoding` に変更される。
 - `deluxemulti=<値>`： `upLATEX` で `japanese-otf` パッケージを `deluxe` オプションおよび `multi` オプション付で使用している (`\UTFK` / `\UTFC` / `\UTFT` が多ウェイトで使える) ときに、日本語以外の CJK 言語のテキストについても多ウェイトの設定を有効にするか否か。
 - `new` / `true`： `TEX Live 2022` 以降で利用可能^{*4}になっている、中国語・韓国語のテキスト用の多ウェイトの論理フォントを使用する。
 - `old`： 本来は `\UTFK` / `\UTFC` / `\UTFT` 入力用である“単純なメトリック”の論理フォントを使用する。約物の周りの空き調整が無効になるという欠点がある。
 - `false` (既定値)： 日本語以外の CJK 言語は多ウェイトにしない。
 - `deluxemulti`： `deluxemulti=new` と同じ。
 - `nodeluxemulti`： `deluxemulti=false` と同じ。
 - `forcedeluxemulti`： `deluxemulti=old` と同じ。
 - `noforcedeluxemulti`： `deluxemulti=false` と同じ。
 - ※ `(no)forcedeluxemulti` は**非推奨**であり、将来に廃止される予定である。
 - `(no)patchutfcmds`： このオプションは何もしない。
 - ※ 1.3 版で試験的に設けられた機能が 1.4 版で削除された。オプション自体も将来に廃止される予定であり、このオプションの指定自体を**非推奨**とする。

3 ユーザ命令

本パッケージの主要機能である CJK フォント切替については Babel の機能を用いて行うことになるため、パッケージ特有の命令は存在しない。

例えば、次のようなパッケージ読込 (一括読込モード) の場合^{*5}：

^{*4} 当該の論理フォントが利用可能でない環境でこの設定を使用すると、`LATEX` または `DVI` ウェアの実行でエラーが発生する。

^{*5} 使用エンジンが `upLATEX` であるか、または `pLATEX` で `bxutf8` 入力エンコーディングを用いているとする。

```
\usepackage[korean,japanese]{pxbabel}
```

本文中で

日本語文書中に`\foreignlanguage{korean}{ 한글 }`がある。

と入力すれば次の出力が得られる。

日本語文書中に 한글がある。

本パッケージが提供する命令には次のものがある。

- `\cjklanguagenamename{(CJK 言語識別子)}`： その CJK 言語に対する Babel 言語名。
※完全展開可能である。
※例えば、パッケージオプションで `schinese=chinesesimp` と指定した場合、`\cjklanguagenamename{schinese}` は「`chinesesimp`」に展開される。
- `\UTFJ{(Unicode 符号値)}`： 常に「本来の `\UTF` 命令」の動作をする。
※本パッケージで `switchfont` が有効で、かつ `japanese-otf` が `multi` 付で読み込まれている場合に定義される。^{*6}この場合、入力言語が日本語以外の CJK 言語のときに `\UTF` 命令の動作が他言語用 (`\UTFK` 等) に切り替わるので、本来の日本語用の `\UTF` 命令が必要なときには `\UTFJ` を利用すればよい。
※詳細は 5.2 節を参照。

4 キャプション言語の固定に関する設定

Babel における言語切替の指定は大別すると、入力言語のみを切り替えるもの (`\foreignlanguage` 命令、`otherlanguage*` 環境) と基底言語を切り替えるもの (`\selectlanguage` 命令、`otherlanguage` 環境) があり、既定動作ではキャプション言語は基底言語に常に追従する (`caption=switch` 指定時と同じ)。

ところが、実際の言語定義を伴わない仮想言語においては、キャプション文字列の設定は行われない。このため、基底言語を仮想言語に切り替えた場合にはキャプション言語が正しくない事態が生じえる。^{*7}これによる混乱を避けるために、`caption` オプションによるキャプション言語の固定が有用かもしれない。

特に、「Babel によるキャプション文字列の変更を抑制して文書クラスで定めた文字列を常に有効にしたい」という場合は、`caption=default` を指定すればよい。文書開始時の基底言語が仮想言語である場合はこちらの使い方が多いと考えられるので、この場合に限って `caption` の既定値を (Babel の既定動作に相当する `switch` ではなく) `default` としている。

※ 0.9 版より前は常に `default` を既定値としていたが、既定言語が仮想言語でない場合は `default` にする必然性が少ないため、Babel の既定動作に合わせる目的で仕様を変更された。ゆえに、`caption` (および `(no)switchcaption`) オプション指定がなくかつ基底言語が仮想言語でない場合の `caption` 設定は、0.9 版より前は `default`、以降は `switch` と異なることになる。

^{*6} ただし、`\UTFJ` 命令は文書本体開始時に定義されることに注意。プリアンブルでは `\UTF` 命令の切替はまだ発生しないので、日本語用には `\UTF` 命令を使えばよい。

^{*7} ただし「文書開始時以外に (入力言語でなく) 基底言語を仮想言語に切り替える」ことをしなければ問題は起こらない。特に、「日本語の文書クラスを用いて、かつ Babel において仮想言語扱いの日本語を文書全体で基底言語にする」場合は、キャプションは文書クラスで定めたものに固定される。多くの場合これは好都合な動作だろう。

※ `bxbase` パッケージの `\fixcaptiopnlanguage` 命令によるキャプション言語固定の設定は、常に本パッケージでの設定より優先される。この仕様は `main` オプションのものと一貫していないが、互換性のためそのままにしている。

5 CJK フォント切替機能の原理

次の2つの仕組みを用意している。

5.1 標準和文フォントの切替 (`upLaTeX` のみ)

`upTeX` は標準で、CJK 言語の `TeX` フォント (和文 TFM) を用意している。本パッケージは `Babel` において CJK 言語の切替処理 (`\extras...` マクロ) の中に、フォントの切替を追加し、言語の選択に応じて適切なフォントが選ばれるようにする。

※標準和文フォントの切替は、内部的には「和文エンコーディング」を切り替えることで実現している。ところが、標準以外の「和文エンコーディング」が指定されている状態では、他パッケージの (和文フォントに関する) 機能に不具合が生じる可能性がある。これは日本語以外の CJK 言語を使う場合にはある程度は仕方がないであろう。ただし、現状の既定では入力言語が日本語である場合にも非標準の「和文エンコーディング」が適用される。^{*8}`stdjaencoding` オプションを指定すると「入力言語が日本語の場合には標準の和文エンコーディングを使う」動作に切り替わるので、不具合が回避できる可能性がある。^{*9}

5.2 `\UTF` 命令に対する調整

`japanese-otf` パッケージの `multi` オプションは韓国語・中国語を扱うためのもので、これを有効にすると `\UTFK` / `\UTFC` / `\UTFT` の各命令^{*10}で韓国語/簡体字中国語/繁体字中国語のフォントを用いて文字を出力することができる。

本パッケージを読み込むと、`Babel` の入力言語が CJK 言語に切り替わった場合に言語に応じて `\UTF` の動作を「本来の `\UTF`」/ `\UTFK` / `\UTFC` / `\UTFT` の何れかに切り替えるようになる。また、「本来の `\UTF`」の動作をする命令 `\UTFJ` が提供されるので、入力言語が日本語以外のときに日本語の文字を入力したい場合には `\UTFJ` を利用すればよい。

この機能は単独ではあまり有用とはいえないが、`pLaTeX` において `bxutf8` 入力エンコーディングと併用して CJK 言語の UTF-8 直接入力を行う時に有用となる。すなわち、直接入力された CJK 文字は `\UTF` に帰着するので、結果的に `upLaTeX` の時と同じように、現在の入力言語用のフォントで出力されることになる。

※ただし、この方策は、「(`pLaTeX` では) JIS X 0208 にある文字はそもそも `\UTF` に変換されないので常に日本語フォントで出力されてしまう」という欠点をもつ。これを解決するには `pxfltsrc` パッケージ等の機能を用いて文書にテキスト変換フィルタを適用するといった追加の処理が必要になる。

^{*8} 標準の横組/縦組の「和文エンコーディング」は `JY2/JT2` であるが、フォント切替を有効にして入力言語を日本語にした場合には `J20/J30` (`jis2004` 有効時には `J24/J34`) が適用される。

^{*9} 最近の `upLaTeX` の標準設定では、実は非標準の「和文エンコーディング」を使うメリットがほぼなくなっている。従って「常に `stdjaencoding` を指定する」という方針でも構わない。現在、`stdjaencoding` の既定値の変更を検討している。

^{*10} `\UTF` と同様に、これらの命令の引数には Unicode 符号値を 16 進表記で指定する。

6 Babel 一括読込モード

0.9 版より前では、Babel 本体 (babel パッケージ) とは別個に読み込む方式のみを採用していた。ところがこの場合、「Babel の本来の機能は不要で pxbabel のフォント切替だけを用いたい」という場合、かなり不自然な指定をする必要があった。例えば、日本語 (japanese) と韓国語 (korean) を切り替えたい (メインは日本語) という場合は次のようになる。

```
\usepackage[english]{babel}
\usepackage[main=japanese]{pxbabel}
```

※ pxbabel を読んで基底言語を japanese にする必要があるのだが、そうするとまず babel を読む必要がありここで何かダミーの言語 (english を用いた) を指定する必要がある。このままでは基底が english になるので、japanese を基底にするためさらに main オプションが必要となる。

そこで、0.9 版において、babel と一括で読み込むことで、babel の使用言語オプション指定の方法を踏襲できるようにした。babel の読込前に pxbabel を読み込むとこの動作になり、babel も一緒に読み込まれる。

```
\usepackage[korean,japanese]{pxbabel}
```

※ オプションは babel のものと同じ理屈で書く。korean と japanese を用い、かつ基底が japanese なのでこの順番になる。

目的が「upTeX の非 BMP 文字対応の日本語フォントが使いたい」だけの場合^{*11}は以下のようにになる。

```
\usepackage[japanese]{pxbabel}
```

※ japanese のみ指定して、当然これを基底とする。

実際には、4 つの CJK 言語のオプションは「常に」出現するので、上の 2 つの指定は等価であるが、一括読込の場合は babel の習慣に合わせて記述するのが解りやすいと思われる。

6.1 複雑な例

少し複雑な例を挙げておく。ドイツ語 (ngerman)、日本語 (japanese)、簡体字中国語 (simpchinese)、繁体字中国語 (tradchinese) を使用し、かつ前の 3 つには言語オプション定義 (*.ldf) があり、基底は日本語で、キャプション言語の固定は行わないとする。

■分離読込モード

```
\usepackage[ngerman,simpchinese,japanese]{babel}
\usepackage[schinese=simpchinese,tchinese=tradchinese,
caption=switch]{pxbabel}
```

※ 分離読込の場合は babel で最後に指定したものが基底言語。

■一括読込モード

^{*11} 現在の upTeX では既定の和文フォントが既に非 BMP 文字に対応しているため、この使い方は不要である。

```
\usepackage[ngerman,schinese=+simpchinese,tchinese=tradchinese,  
+japanese]{pxbabel}
```

※非仮想言語の `japanese` が基底なので、`caption` は `switch` が既定値になる。なお、非 CJK 言語 (`ngerman`) には決して + を付けないことに注意。