

Ω T_EX と OTF パッケージ

宮崎 泉

概要

文字空間を 16bit に拡張して多言語化に正面から取り組んだΩは、OTP という枠組みを用意することで柔軟な拡張性を持つ。この OTP を利用して、CHISE の多漢字環境を印刷する方法を説明するとともに、漢字以外の文字も含めた多スクリプト環境の印刷の可能性を紹介する。また、多漢字環境だけを必要とする人のために、pT_EX と OTF パッケージを利用した、より日本語の組版に適した方法も紹介したい。

はじめに

- 多言語文書の印刷
- T_EX 日本語化の歴史
 - NTT J_TE_X, pT_EX

1 UTF-8 文書の作成

入力の方法

- OS が用意するもの
- アプリケーションが用意するもの
 - ex.) XEmacs/CHISE

XEmacs/CHISE

- 入手方法とインストールの方法
 - <http://www.kanji.zinbun.kyoto-u.ac.jp/projects/chise/xemacs/>
Debian GNU/Linux と Mac OS X にインストールする方法の説明がある。
- 使い方
 - `ids-find-chars-including-components`
 - `ids-find-chars-covered-by-components`

2 Ω/CHISE による多漢字文書の印刷

■Ω/CHISE とは? TeX を多言語化するために開発されたΩを使って、UTF-8 の文書、あるいは XEmacs/CHISE で作成された UTF-8 互換形式の文書を印刷するためのΩパッケージ

■特長

- XEmacs/CHISE の内部コードを利用できる。
- IDS (*Ideographic Description Sequence*) を使った文字の表記が出来る。(矢口など)
- 既存の文字セットにある字形はそれを利用する。
- 既存の文字セットにない字形は、Kage サーバを利用して文字合成を行う。

■入手方法とインストールの方法

- <http://www.kanji.zinbun.kyoto-u.ac.jp/projects/chise/omega/INSTALL.txt.ja> あるいは [INSTALL-MacOS-X.txt.ja](#) を参照。

■Ω/CHISE の使用法

- パッケージオプション (プリアンブルでの指定)
 1. 入力コードの指定 [必須]
(`utf8mcs`, `utf8gb`, `utf8cns`, `utf8jis`, `utf8ks`)
 2. 本文中のベースとなるフォントの指定 [オプション]
(`gbfont`, `cnsfont`, `jisfont`, `ksfont`, `multifont`)
 3. Kage を利用するかどうか [オプション]
(`kage`)

例)

```
\usepackage[utf8mcs,multifont,kage]{chise}
```

- 環境を利用したベースとなるフォントの切り替え
- `GBfont`, `CNSfont`, `JISfont`, `KSfont`, `Multifont`

例)

```
\begin{JISfont}  
日本語  
\end{JISfont}
```

■Ω/CHISE の仕組み

- Ω TP (Ω Translation Process) (1)
 - CHISE DB を参照するために CHISE 内部コードに変換
 - IDS をひとつの文字として扱い、可能であれば CHISE 内部コードへ変換
 - CHISE 内部コードと Adobe CID を照合。CID が存在すればそのフォントを利用
 - ターゲットとするフォントに文字がない場合 IDS に変換
- Ω TP (2)
 - KAGE サーバの呼び出し
 - PostScript Type 1 フォントの生成 (外部プログラム t1asm, fontforge の呼び出し)

■Ω/CHISE の問題点

- Ωそのものが開発途上にあるため、Ωと整合性がとれず動作しなくなることがある。
- Kage が発展途上にあり、合成出来ない文字も多い。
- 組版エンジンが日本の組版ルールに対応したものではないため、OTP を利用して日本の組版ルールを実現しなければならないが、現時点のΩ/CHISE では不完全。

3 Ωの応用: 多スクリプト文書の印刷

- Ω TP によって他のスクリプトに対応付けることで、多スクリプト文書の印刷が可能になる。
- Ω TP を切り替えることで、ラテン文字での転写と本来のスクリプトでの印刷を切り替えることも出来る。
 - デーヴァナーガリー ⇔ ラテン文字の転写
 - チベット文字 ⇔ ラテン文字の転写 など

4 chise2otf

■chise2otf とは？ UTF-8 で書かれた TeX ファイルあるいは XEmacs/CHISE で作成された TeX ファイルを、OTF パッケージのマクロを使った TeX ファイルに変換する Perl スクリプト

- 長所
- 組版に pLaTeX を利用するので、日本語の組版に適している。
 - OTF パッケージのマクロを意識することなく、OTF パッケージを利用できる。
- 短所
- chise2otf を使って変換する手間が増える
 - Ωのように環境によって、ベースとなるフォントを切り替えることは出来ない。

OTF パッケージ

- 入手方法とインストールの方法
 - <http://psitau.at.infoseek.co.jp/otf.html>
- マニュアル
 - <http://psitau.at.infoseek.co.jp/PDF/otfmanual.pdf>
- 関連文書 (Adobe 社による CID の解説)
 - <http://partners.adobe.com/asn/developer/pdfs/tn/5078.Adobe-Japan1-6.pdf>

例)

```
\UTF{76F4}, \UTFC{76F4}, \UTFK{76F4}, \UTFT{76F4}, \UTFM{76F4}  
\CID{3034}
```

出力 直, 直, 直, 直, 直
直

■ 使用法

chise2otf [-i 入力コード] [-o 漢字優先順位] [-klru] ファイル名

- i 入力コード (デフォルト: ucs@mcs)
ucs@mcs, ucs@cns, ucs@gb, ucs@jis, ucs@ks
- o 漢字優先順位: (デフォルト: j)
c: CNS, g: GB, j: JIS, k: KS, G: GT 2000 書体
m: OTF パッケージのマクロ\UTFM を使う。組み合わせて使うことも出来る (例) jGcgkm)
- k Kage サーバを利用して文字合成を行う
- l ucs@jis 環境のラテン文字をそのまま使う
- r 下点付き r と l を下丸付き r と l に置換する
- u Unification を許す

■GT 2000 書体 東京大学多言語処理研究会によって開発された 6 万 7 千字を超える漢字の TrueType フォント

入手先 <http://www.l.u-tokyo.ac.jp/GT/>