glc-utf8.sed 変換内容

松浦高志

2006 年 12 月 12 日

拙作の glc-utf8.sed スクリプトを用いると、UTF-8（Unicode）で書かれた文字や記号のうち、EUC で表せないものを、必要に応じて種々のパッケージを用いて LaTeX の記法に変換するものである。UTF-8 で規定されているすべての文字を網羅するのは非常に大変なので、西洋古典学などでよく使われるものだけを選んでいる。

変換は sed コマンドを用いて次のように行う。UNIX で処理することを考え、改行コードは UNIX 標準の LF にしてある。

```
$ sed -f glc-utf8.sed input.txt >output.txt
```

いくつかパッケージが必要になる場合がある。この文書で使われているパッケージの一覧を次に示す。

```
\usepackage[T1]{fontenc}
\usepackage{textcomp}
\usepackage{pxfonts}
\usepackage[greek,english]{babel}
\usepackage{teubner}
\usepackage{pifont}
\usepackage{okumacro}
```

pxfonts が指定されているのは、フォントによっては表示できない文字があるからである。

このスクリプトによって変換される文字は次の通りである。

ÆŒæœłŁ¿ ¡ ıøØÅå

ÞþÐð

À Á Â Ã Ä Ā Ą Ă

Ç Ć Č Ĉ Ċ

Ď

È É Ê Ë Ē Ę Ě Ĕ Ė

Ĝ Ğ Ġ

Ĥ

Ì Í Î Ï Ī Ĩ Ĭ Ị İ

Ĵ

Ķ

Ĺ
Ñ Ń Ň
Ò Ó Ô Õ Ö Ō Ő Ŏ
Ŕ Ř
Ś Š Ş T̂ Ť Ţ
Ù Ú Û Ü Ū Ů Ű Ŭ Ũ Ų
Ŵ
Ý Ŷ Ÿ
Ź Ž Ż
à á â ã ä å ā ą ă
ç ć č c̃ ċ
è é ê ë ē ę ě ĕ ė
ĝ ğ ġ
ĥ
ìíîïĩīĭ į ı
ĵ
ĺ
ñ ń ň
ò ó ô õ ö ō ő ŏ
ŕ ř
ś š ş t̂ ţ ť ú û ü ū ů ű ŭ ũ ų
ŵ
ý ÿ ŷ
ž ż ź
ϹF
ϙ
ƛϟ
ϛϝϙϡϟ
Ά Ἀ Ἁ Ὰ Ά Ᾰ Ᾱ ᾼ Ἂ Ἃ Ἄ Ἅ Ἆ Ἇ ᾈ ᾉ ᾊ ᾋ ᾌ ᾍ ᾎ ᾏ
Έ Ἐ Ἑ Ὲ Έ Ἒ Ἓ Ἔ Ἕ
Ή Ἠ Ἡ Ὴ Ή ῌ Ἢ Ἣ Ἤ Ἥ Ἦ Ἧ ᾘ ᾙ ᾚ ᾛ ᾜ ᾝ ᾞ ᾟ
Ί Ἰ Ἱ Ὶ Ί Ῐ Ῑ Ϊ Ἲ Ἳ Ἴ Ἵ Ἶ Ἷ
Ό Ὀ Ὁ Ὸ Ό Ὂ Ὃ Ὄ Ὅ
Ῥ
Ύ Ϋ Ὑ Ὺ Ύ Ῠ Ῡ Ὓ Ὕ Ὗ
Ώ Ὠ Ὡ Ὼ Ώ ῼ Ὢ Ὣ Ὤ Ὥ Ὦ Ὧ ᾨ ᾩ ᾬ ᾭ ᾪ ᾫ ᾮ ᾯ
άἀἁὰάᾰ ᾱ ᾶᾳἂἃἄἅἆἇᾲᾴᾷᾀᾁᾂᾃᾄᾅᾆᾇ
έἐἑὲέἒἓἔἕ
ήἠἡὴήῆῃἢἣἤἥἦἧᾐᾑῂῄῇᾒᾓᾔᾕᾖᾗ
ίἰἱὶί ῑ ῖϊΐἲἳἴἵἶἷῒΐῗ

ȯȯȯȯȯȯȯȯȯ

ṗṗṷṷṷṷṷ ṷ ṷṷṷṷṷṷṷṷṷṷṷ

ẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇẇ

〈〉‹›«»

. . . – — - ~•“ ”

‘ “ ’ ‰ ‰ooo§¶†‡¦ˆ˜ao–o˜˜˜.˝

©®℗SMTM№µΩ°C

½¼¾

0 1 2 3 4 5 6 7 8 9

←→↑↓

¢£¥฿€¤₦₩

±×∗÷√

✂ ✁ ✃ ✄ ☎ ✆ ✇ ✈ ✉ ☛ ☞ ✌ ✍ ✎ ✏ ✐ ✑ ✒ ✓ ✔ ✕ ✖ ✗ ✘ ✙ ✚ ✛ ✜ ✝ ✞ ✟ ✠ ✡ ✢ ✣ ✤ ✥ ✦ ✧ ☆ ✪ ★ ✫ ✬ ✭ ✮ ✯ ✰ ✱ ✲ ✳ ✴ ✵ ✶ ✷ ✸ ✹ ✺ ✻ ✼ ✽ ✾ ✿ ❀ ❁ ❂ ❃ ❄ ❅ ❆ ❇ ❈ ❉ ❊ ❍

❍ ❑ ❒ ❑ ❒ ▲ ▼ ◆ ❖ ◗ ❙ ❚ ❛ ❜ ❝ ❞ ❡ ❢ ❣ ❤ ❥ ❦ ❧ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ❶ ❷ ❸ ❹ ❺ ❻ ❼ ❽ ❾ ❿ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ❶ ❷ ❸ ❹ ❺ ➿ ❼ ❽ ❾ ❿ ➔

↔ ↕ ↘ ↙ ↗ ↖ ➔ → ➡ ➡ ➡ ➢ ➢ ➢ ➡ ➡ ➡ ➡ ➡ ➪ ➪ ➪ ➪ ➪ ➪

➪ ➲ ➳ ➴ ➵ ➶ ➷ ➸ ➹ ➺ ➻ ➼ ➽ ➾