



HAL
open science

Asian TeX-like typeset engines

Jean-Michel Hufflen

► **To cite this version:**

Jean-Michel Hufflen. Asian TeX-like typeset engines. BachoTeX 2008 Conference, 2008, Poland. pp.129–131. hal-00661894

HAL Id: hal-00661894

<https://hal.science/hal-00661894>

Submitted on 20 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asian T_EX-Like Typeset Engines*

Jean-Michel HUFFLEN

LIFC (EA CNRS 4157)

University of Franche-Comté

16, route de Gray

25030 BESANÇON CEDEX

FRANCE

hufflen@lifc.univ-fcomte.fr

<http://lifc.univ-fcomte.fr/~hufflen>

Abstract

In order to extend MIBIB_{T_EX} to languages of the Far East, we are experiencing _{T_EX} engines for them — e.g., p_{T_EX} — after attending the first Asian _{T_EX} conference. We propose a demonstration of that.

Keywords Asian _{T_EX} engines, p_{T_EX}, CJK package, kotex package.

Streszczenie

Przy rozszerzaniu MIBIB_{T_EX}-a do pracy z językami Dalekiego Wschodu, dzięki udziałowi w 1-szej konferencji Asia _{T_EX}, mogliśmy zająć się odpowiednimi silnikami _{T_EX}-owymi, np. p_{T_EX}-em. Zademonstrujemy wyniki.

Słowa kluczowe warianty _{T_EX}a dla języków dalekowschodnich, p_{T_EX}, pakiet CJK, pakiet kotex.

0 Introduction

In January 2008, the first Asian _{T_EX} conference was held at Kongju, in South Korea. In fact, presentations focused on languages used throughout the Far East, since most people attending at this conference came from Korea, China, Japan, and Vietnam. Concerning us, we are interested in enlarging MIBIB_{T_EX}¹ [5], our reimplementations of BIB_{T_EX} [14], to Asian languages. Going to this conference allowed us to get more familiar with the packages and engines used for these languages, for which only a little documentation is written in English.

Roughly speaking, the present situation about Asian languages is the same than at the beginning of the 1990's for European languages: some *ad hoc* packages — french [3], german [16], polish [2, § F.7] — were already being developed, but without global organisation. Each package allowed users to write both in their own language and English. The babel package was only announced in *The L^AT_EX Companion*'s first edition [4, Ch. 9]; it is now fully working and described in *The L^AT_EX Companion*'s sec-

* Title in Polish: *T_EX-owe narzędzia składu dla języków Dalekiego Wschodu*

¹ MultiLingual BIB_{T_EX}. This program has been used to build the current article's bibliography. Asian person names are put down according to their tradition, that is, the last name comes before the first name.

ond edition [12, Ch. 9]. Let us recall that this package aims to process all the languages it knows in a homogeneous way². However, the present distribution does not include support for Far East languages by means of this package, although some tools presented hereafter — the packages CJK and kotex — can be used in conjunction with it.

Section 1 gives a short overview of the CJK package. Then we cite two *ad hoc* ways to write in Korean and Japanese in Sections 2 and 3. In addition to these packages and engines, let us just cite Vn_{T_EX} for the Vietnamese language, including Vietnamese fonts³, support for L^AT_EX and Plain _{T_EX} [18]. The definitions suitable for L^AT_EX are usable by means of a vietnam package or a vietnam option of the babel package.

1 The CJK package

‘CJK’ is a collective term for ‘Chinese, Japanese, Korean’, since these writing systems completely or partly use Chinese characters. ‘CJK’ characters are implemented in an area of Unicode's basic multilingual plane [19], inside the range U+3000–U+9FA5.

² In fact, it encompasses most of the features provided by packages such as french, german, polish.

³ Vietnamese is written with Latin letters, with many accented letters. Some letters can have two accents, the total number of accented letters in Vietnamese being 134 [18].

In fact, the Chinese and Japanese languages use ideograms, whereas modern Korean is written using the Hangul alphabet, organised into syllabic blocks. Each block is a combination of consonants and vowels and notes a syllable down. Traditionally, Chinese and Korean are written horizontally. Often the Japanese language uses vertical writing according to its tradition, but the horizontal sense is used for technical writing.

The CJK package [10] implements several encoding schemes to reach the characters belonging to the CJK area and some basic typographic rules related to these three languages. It can be used in conjunction with the `babel` package. Its limitation: that is preferable for it if a non-CJK language is the main language⁴. This package has been improved into a `CJKutf8` package [11], dealing with UTF-8⁵ texts more easily, and a version derived from it — the `xCJK` package — is usable with the `XYTEX` typeset engine⁶ [6].

2 The `koYTEX` package

The steps of developing a Hangul `TEX` are related in Korean in [1]. It seems that Korean people presently use the tools of the `koYTEX` distribution [8], including the `kotex` package and the `halpha` bibliography style, usable with `BIBYTEX`. This `kotex` package is able to deal with UTF-8 texts. Text fragments — including texts put within commands' arguments — are supposed to be in Korean (resp. English) if Asian (resp. Latin) characters are used. Besides, it can be used in conjunction with the `babel` package.

The `LATEX` source text of a very short example is given at Figure 1. This complete example — as the `ajt-sample.tex` file — belongs to the document class files for the *Asian Journal of TEX*⁷, we have just replaced the fragments using Korean characters by occurrences of '`...`'. When this text is processed by `LATEX`, the `korean` option of the `ajt` document class causes the `kotex` package to be loaded, and the result is given at Figure 2.

⁴ As abovementioned, this CJK package mainly aims to access CJK characters and assemble them into fragments. But — as an example — for the whole of a text written using one of CJK languages, the spacing between two baselines should be wider, that is, the value of the `\baselineskip` command [7, Ch. 12] should be enlarged. So do the `kotex` package (cf. § 2) and the `pYTEX` engine (cf. § 3).

⁵ **Unicode Transformation Format**. The UTF-8 encoding is probably the most used now for texts including non-Latin characters. See [19] for more details.

⁶ `XYTEX` is a typesetting system based on a merger of `TEX`'s system with `Unicode` and modern font technologies.

⁷ You can download all these files from <http://ftp.ktug.or.kr/pub/ktug/ajt/>.

```
%% Select pdfYTEX or your favorite DVI driver.
\documentclass[korean]{ajt} % pdfYTEX (default).
%%

%% Mandatory article metadata.
\title[How to prepare a document with AJT
class]{AJT (...)}
\author[Gildong Hong]{(...)}
\address{(...)}
\email{gdhong@ktug.or.kr}

\abstract{(...)} AJT (...)(abstract)(...)
preamble(...)} AJT (...)\PracTeX(...)}
%%

\begin{document}

\maketitle
%%

\section{(...)}
(...)

\subsection{(...)}
(...)

\subsection{(...)}
(...)\verb+\[+ (...)\verb+\]+ (...)

\begin{thebibliography}{9}
\bibitem[1]{...} ...

\bibitem[2]{...} ...

\bibitem[4]{...} ...

\bibitem[5]{...} ...
\end{thebibliography}
%%

\end{document}
```

Figure 1: Example using the `kotex` package: source text.

3 Japanese `TEX`

`pYTEX` [13] is a typeset engine based on `TEX` and implementing formatting rules for Japanese documents. In particular, it deals with the two possible senses of writing — horizontal or vertical —, provides some *ad hoc* document classes `jsarticle`, `jsbook`, `jreport`. Its new version, `upYTEX` [17] may be viewed as 'a new `pYTEX`', dealing with `Unicode` and supporting the `babel` package. An integration of `upYTEX` and Ω^8 [15], `upOmega`, is also planned.

To end, let us mention the development of a Japanese `TEX` environment for `Cygwin`, usable under `Windows` [9].

⁸ Ω is an extension of `TEX` aiming to improve multilingual typesetting.

For submission to The Asian Journal of \TeX
Draft of March 10, 2008

AJT 클래스 사용에 대하여

How to prepare a document with AJT class

홍길동 Gildong Hong
덕대학교 대학학 석학과 gdhong@ktug.or.kr

ABSTRACT 다른 클래스들과 달리 AJT 클래스는 초록(abstract)이 preamble에 있어야만 한다. 이것은 AJT 클래스의 모태가 되는 Prac \TeX 이 그렇게 디자인되어있기 때문이다.

1 첫째 절

첫째 절은 다음과 같은 부절들로 이루어져있다.

1.1 첫째 부절

첫째 부절은 다음과 같은 내용들을 가진다. 수식 다음에 공백을 넣지 않아도 자동으로 공백이 들어간다. 예를 들어 x 다음에는 공백이 들어가 있지만 x 는 공백없이 입력되었다.

1.2 둘째 부절

둘째 부절은 다음과 같은 내용들을 가진다. 디스플레이(display) 수식을 사용할 때에는 \$\$는 절대 사용하지 말고 대신 \[및 \] 를 사용하는 것이 바람직하다.

참고 문헌

1. Donald E. Knuth, *The \TeX book*, Addison-Wesley, 1986.
2. Jonathan Fine, *Some Basic Control Macros for \TeX* , TUGboat 13 (1992), no. 1, 75–83. <https://www.tug.org/TUGboat/Articles/tb13-1/tb34fine.pdf>
3. Victor Eijkhout, *\TeX by Topic, A \TeX nician's Reference*, Addison-Wesley, 1992. <http://www.cs.utk.edu/~eijkhout/textbytopic-a4.pdf>
4. <http://en.wikipedia.org/wiki/TeX>

Figure 2: Example using the kotex package: result.

4 Acknowledgements

Many thanks to Jerzy B. Ludwichowski, who has written the Polish translation of the abstract. Tomasz Przechlewski has helped translate keywords, thanks to him, too.

References

- [1] CHO Jin-Hwan: “The Passage of Hangeul \TeX and ko \TeX ”. *The Asian Journal of \TeX* , Vol. 1, no. 2, pp. 113–121. October 2007.
- [2] Antoni DILLER: *L^A \TeX wiersz po wierszu*. Wydawnictwo Helio, Gliwice. Polish translation of *L^A \TeX Line by Line* with an additional annex by Jan Jelowicki. 2001.
- [3] Bernard GAULLE : *Notice d’utilisation du style french multilingue pour L^A \TeX* . Version 4.05b. Février 1999. Joint à la distribution de L^A \TeX .
- [4] Michel GOOSSENS, Frank MITTELBACH and Alexander SAMARIN: *The L^A \TeX Companion*. 1st edition. Addison-Wesley Publishing Company, Reading, Massachusetts. 1994.
- [5] Jean-Michel HUFFLEN: “MIBIB \TeX : Reporting the Experience”. *TUGboat*, Vol. 29, no. 1, pp. 157–162. EuroBacho \TeX 2007 proceedings. 2007.
- [6] Jonathan KEW: “X_Y \TeX in \TeX Live and beyond”. *TUGboat*, Vol. 29, no. 1, pp. 146–150. EuroBacho \TeX 2007 proceedings. 2007.
- [7] Donald Ervin KNUTH: *Computers & Typesetting. Vol. A: The \TeX book*. Addison-Wesley Publishing Company, Reading, Massachusetts. 1984.
- [8] Korean \TeX Society, <http://project.ktug.or.kr/ko.TeX>: *Korean \TeX* . October 2007.
- [9] KUROKI Yusuke: “Japanese \TeX Environment for Cygwin”. Slides for the 1st Asian \TeX Conference. Kongju National University, South Korea. January 2008.
- [10] Werner LEMBERG: “The CJK Package: Multilingual Support beyond babel”. *TUGboat*, Vol. 18, no. 3, pp. 214–224. September 1997.
- [11] Werner LEMBERG: “Unicode Support for the CJK Package”. Hand-out version for the 1st Asian \TeX conference. Kongju National University, South Korea. January 2008.
- [12] Frank MITTELBACH and Michel GOOSSENS, with Johannes BRAAMS, David CARLISLE, Chris A. ROWLEY, Christine DETIG and Joachim SCHROD: *The L^A \TeX Companion*. 2nd edition. Addison-Wesley Publishing Company, Reading, Massachusetts. August 2004.
- [13] OKUMURA Haruhiko: “Japanese \TeX . Past, Present, and Future”. Slides for the 1st Asian \TeX Conference. Kongju National University, South Korea. January 2008.
- [14] Oren PATASHNIK: *BIB \TeX ing*. February 1988. Part of the BIB \TeX distribution.
- [15] John PLAICE and Yannis HARALAMBOUS: “The latest developments in Ω ”. *TUGboat*, Vol. 17, no. 2, pp. 181–183. June 1996.
- [16] Bernd RAICHLE: *Die Makropakete „german“ und „ngerman“ für L^A \TeX 2 ϵ , L^A \TeX 2.09, Plain- \TeX and andere darauf Basierende Formate*. Version 2.5. Juli 1998. Im Software L^A \TeX .
- [17] TANAKA Takuji: “Short Introduction to up \TeX —Unicode p \TeX ”. Slides for the 1st Asian \TeX Conference. Kongju National University, South Korea. January 2008.
- [18] THÀNH HÀn THÉ: “Typesetting Vietnamese with Vn \TeX (and with the \TeX Gyre fonts too!)”. *TUGboat*, Vol. 29, no. 1, pp. 95–100. EuroBacho \TeX 2007 proceedings. 2007.
- [19] THE UNICODE CONSORTIUM: *The Unicode Standard Version 5.0*. Addison-Wesley. November 2006.