X∃TEX: blending the best of Mac OS X and TEX

Jonathan Kew

SIL International

June 18, 2007

History

- April 2004: first public release of X∃TEX 0.3 (for Mac OS X only)
    - built-in support for Unicode
    - access to all fonts installed on the user's computer
    - AAT (Apple Advanced Typography) for typographic features
    - graphics support through QuickTime
- February 2005: X∃TEX reaches version 0.9
    - OpenType layout features
    - mature LATEX support packages
- April 2006 (BachoTEX): X∃TEX for Linux (first public release)
- June 2006: X∃TEX built for Windows by Akira Kakuto
- February 2007: TEX Live 2007 includes X∃TEX 0.996 for all platforms
    - many thanks to Karl Berry, and all TL builders and testers!

Key features of the X∄TEX system
included in current MacTEX, TEX Live, and gwTEX distributions

- Unicode supported as the standard text encoding
- most macro packages (except input encoding and font support packages based on non-Unicode encodings) work unchanged
  - important packages such as `graphics`, `color`, `geometry`, `crop`, `hyperref` recognize X∄TEX automatically
- `xu-hyphen` wrappers load hyphenation patterns in Unicode
- use of OpenType, TrueType, PostScript fonts with no TEX-related setup procedure
- OpenType layout features (ligatures, swashes, alternate glyphs, dynamic accent attachment, etc.)
- extended Latin and complex non-Latin scripts through industry-standard Unicode encoding and fonts

Русский हिन्दी 汉语 نستعليق اردو 한국어 Ελληνικά עברית 日本語

Typesetting Unicode text with X∄TEX

- accented characters (many more than in any legacy codepage)

```
\font\txt="Charis SIL" at 14pt \txt
Hej Slované, ještě naše slovanská řeč žije.
Óðinn átti tvá brœðr. Hét annarr Vé, en annarr Vílir.
Dünyayı verelim çocuklara hiç değilse bir günlüğüne.
Kuř béga Šešùpė, kuř Nēmunas tēka, taĩ mūsų tėvỹnė, ...
```

Hej Slované, ještě naše slovanská řeč žije.

Óðinn átti tvá brœðr. Hét annarr Vé, en annarr Vílir.

Dünyayı verelim çocuklara hiç değilse bir günlüğüne.

Kuř béga Šešùpė, kuř Nēmunas tēka, taĩ mūsų tėvỹnė, gražì Lietuvà.

Far Eastern scripts (Chinese, Japanese, Korean)

- they're just more characters; no special effort required

```
\font\han="STSong" at 16pt
\font\rom="Gentium" at 8pt

\def\hc#1#2{\vbox{\hbox{\han #1}
   \hbox{\kern10pt\rom #2}}}

\vbox{\hc{書く}{ka-ku}
   \hc{最も}{motto-mo}
   \hc{最後}{sai-go}
   \hc{働く}{hatara-ku}
   \hc{海}{umi}}
```

書く
ka-ku

最も
motto-mo

最後
sai-go

働く
hatara-ku

海
umi

Vertical text with XƎTEX

```
\font\body="STKaiti:vertical" at 7pt \body
\font\bold="STHeiti:vertical" at 7pt
\font\title="STHeiti:vertical" at 10pt
\centerline{\title 三  国  演  义} \medskip
\centerline{\bold 〔明〕罗贯中} \smallskip
\leftline{词曰：}
滚滚长江东逝水，浪花淘尽英雄。是非成败转头空：青山依旧
在，几度夕阳红。
白发渔樵江渚上，惯看秋月春风。一壶浊酒喜相逢：古今多少
事，都付笑谈中。
\medskip
\centerline{\bold 第一回} \smallskip
\centerline{\bold 宴桃园豪杰三结义        斩黄巾英雄
首立功}\medskip
话说天下大势，分久必合，合久必分：周末七国分争，并入于
秦；及秦灭之后，
楚、汉分争，又并入于汉；汉朝自高祖斩白蛇而起义，一统天
下，后来光武中兴
，传至献帝，遂分为三国。推其致乱之由，殆始于桓、灵二
帝。桓帝禁锢善类，
崇信宦官。及桓帝崩，灵帝即位，大将军窦武、太傅陈蕃，共
相辅佐；时有宦官
曹节等弄权，窦武、陈蕃谋诛之，机事不密，反为所害，中涓
自此愈横。\par
```

三 国 演 义

〔明〕罗贯中

词曰：滚滚长江东逝水，浪花淘尽英雄。是非成败转头空：青山依旧在，几度夕阳红。白发渔樵江渚上，惯看秋月春风。一壶浊酒喜相逢：古今多少事，都付笑谈中。

第一回

宴桃园豪杰三结义　斩黄巾英雄首立功

话说天下大势，分久必合，合久必分：周末七国分争，又并入于秦；及秦灭之后，楚、汉分争，又并入于汉；汉朝自高祖斩白蛇而起义，一统天下，后来光武中兴。传至桓、灵二帝。桓帝禁锢善类，崇信宦官。及桓帝崩，灵帝即位，大将军窦武、太傅陈蕃，共相辅佐；时有宦官曹节等弄权，窦武、陈蕃谋诛之，机事不密，反为所害，中涓自此愈横。

Demos of X∃TEX font support

- multi-script support: Latin, Devanagari, Arabic, Chinese

$$Hello\ world!$$

नमस्ते दुन्या!

السلام عليكم!

你好世界!

```
\fontspec{Zapfino} Hello world! \par
\fontspec{Devanagari MT} नमस्ते दुन्या! \par
\fontspec{Geeza Pro} السلام عليكم! \par
\fontspec{STKaiti} 你好世界! \par
```

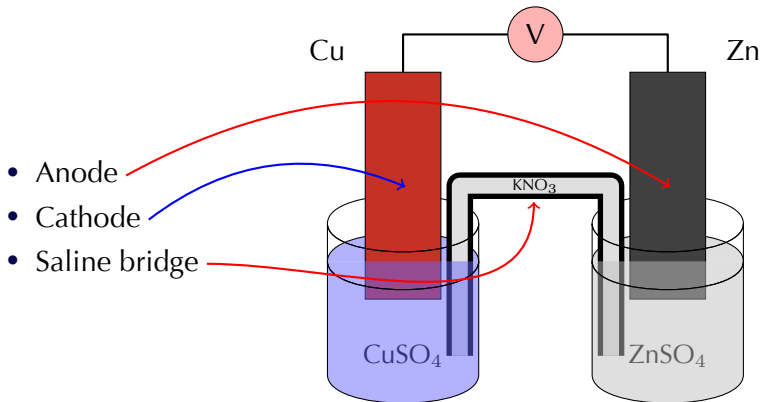- installing an OpenType font, and using it in X∃LATEX via fontspec.sty

Graphics support

- recommended formats: PDF, PNG, JPG
- on Mac OS X, many other formats also supported (everything QuickTime can handle)
- LATEX `graphicx` package automatically supports X∃TEX
- underlying primitives `\XeTeXpdffile`, `\XeTeXpicfile`
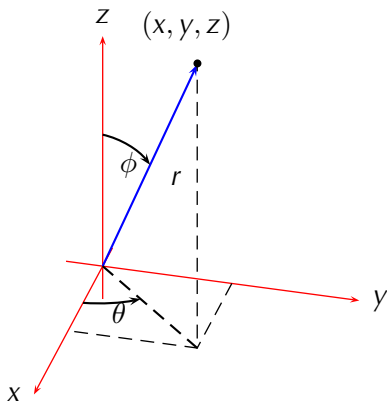
Drawing packages such as TikZ[1]

- works because `pgf` supports the `dvipdfm(x)` driver

Support for PSTricks[2]

- extension to `xdvipdfmx` driver (Miyata Shigeru)

- demo: `03-3d12.tex`

[2]Example from http://www.tug.org/PSTricks/

Inter-character token lists

- uses new `\XeTeXcharclass` property of all Unicode characters
- inserts arbitrary token lists between adjacent characters, based on their class
- possible applications
    - sophisticated CJK spacing
    - automatic font switching between scripts
    - letter-spacing for special effects
- demo: `04-interchar`

Graphite font rendering

- Graphite: open-source library for font layout features
  - support for complex scripts, minority languages, extensions to Unicode
  - all glyph layout features controlled by the font
  - can provide script behavior for Private Use characters, allowing implementation before standardization
- demo: `05-graphite`

Extended Unicode math support

- traditional TEX math formatting is heavily dependent on `tfm` data and macros that have detailed knowledge of multiple special font encodings
- number of math families increased from 16 to 256
- extended versions of TEX's `\mathchar`, `\delcode` etc. supporting full Unicode range (including supplementary-plane characters)
- support for `MATH` table in OpenType fonts
  - provides glyph metrics for math mode
  - variable-size operators
  - extensible delimiters, radicals, etc.

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right)$$

- demo: 06-cambria-math

Under consideration for future releases
these are possibilities, not promises!

- Unicode-based Babel replacement
  - document language (fixed strings)
  - hyphenation language (patterns)
  - typographic language or locale (quotes, spacing, other conventions…)
  - associating font with script or language when required
- extended hyphenation options
  - extend `\hyphenchar` to `\prehyphentext` and `\posthyphentext` to permit insertion of Unicode controls, post-break character, etc.
  - question: associate this with the font (like `\hyphenchar`) or with the `\language` setting?

Under consideration (continued)
these are possibilities, not promises!

- pdfTEX-like microtypographic refinements
  - hanging punctuation, margin kerning
  - font expansion
- adopt some key features from Omega
  - full text directionality model
  - local left/right box
- tighter integration with TeXShop and other front-ends
  - automatic generation of `pdfsync` information
  - linkback support to open graphics files directly from the preview
  - MathML formatting engine (with Unicode math fonts) suitable for use within Safari, etc.
  - …what else?

Thanks

Jin-Hwan Cho	for making a cross-platform output driver possible
Akira Kakuto	for building X∃TEX on Windows
Ross Moore	for LATEX compatibility packages
Will Robertson	for fontspec and additional macro packages
Miyata Shigeru	for assistance with vertical text issues, and PSTricks support
Karl Berry	for supporting the integration into TEX Live
the user groups	for opportunities to meet the community, and for supporting my participation
many others	especially on the X∃TEX mailing list for ideas, bug reports, encouragement, etc.
Don Knuth	who started it all by giving us TEX

Questions… and answers?

- Contact information
  - `mailto:jonathan_kew@sil.org`
- X∃TEX web site and mailing list
  - `http://scripts.sil.org/xetex`
  - `http://tug.org/mailman/listinfo/xetex`
- Source repository for ongoing development
  - `http://scripts.sil.org/svn-public/xetex/TRUNK`

የኢኮድ ምንድን ነው? ما هي الشفرة الموحدة "يونكود"؟ 什麼是Unicode(統一碼/標準萬國碼)? Što je Unicode? ᲠᲐ ᲐᲠᲘᲡ �უᲜᲘᲙᲝᲓᲘ? Tí εἶναι τὸ Unicode; מה זה יוניקוד? यूनिकोड क्या है? Hvað er Unicode? ユニコードとは何か? 유니코드에 대해? يونيکُد چيست؟ Что такое Unicode? Unicodeคืออะไร? የኢኮድ አንታይ ኢዩ?