

Pan-Unicode Fonts

Text Layout Summit 2007
Glasgow, July 4-6

Ben Laenen, DejaVu Fonts

Introduction

Feature request last Friday for DejaVu:

Request for Khmer characters U+1780-17DD, 17E0-17E9, 17F0-17F9:

[...]

The DejaVu Sans Basic Latin characters are more readable than those in the KhmerOS fonts. And DejaVu fonts contain IPA phonetic characters that KhmerOS fonts lack.

Thus it is currently impossible to put English, Khmer and pronunciations in the same text to create a dictionary, except in two word-processing programs that can handle Khmer ZWSP `_and_` use more than one font in the same text.

Would it be possible to adapt one of the KhmerOS fonts (or more than one of them, for Sans, Serif and Mono) and provide Khmer characters for the 114 characters?

Until more programs are capable of displaying different Unicode sections with different fonts -- a la Firefox -- it is of critical importance to get all the characters into a single font.

[...]

Introduction

- Current policy in DejaVu: if someone designs the glyphs, they'll get in the fonts.
- Issue raised: is it really in the scope of DejaVu to have all these kinds of scripts?
- Is there a need for Pan-Unicode fonts?

Why pan-Unicode fonts

- Pan-Unicode fonts keep popping up... First the FreeFonts, also DejaVu now
- See request for Khmer: users want to have nice Latin letters next to script X when they typeset it
- Style issue: finding two matching fonts of different scripts is very hard sometimes
- Many scripts also use common punctuation and digits; those glyphs are in both fonts preferably the same

So...

- Given the fact these issues are valid, DejaVu keeps getting new scripts
- We do have special LGC version with only Latin, Greek and Cyrillic.

Discussion about it

- There has been a lot of discussion about it:
- Don't put it all in one font
- Stick to what you know best
- Metrics of difficult scripts are incompatible

Stick to what you know best

- Often people making glyphs for a script aren't familiar with it
- Will evidently result in a design that has some problems
- e.g.: Code2000 has a lot of scripts that are unusable

On the other hand...

- The project is alive: glyphs will get better over time
- User interaction: if glyphs aren't good, users *will* complain
- Reviews by native speakers during initial design phase (if possible, it's sometimes very hard to find suitable people)
- Given the fact it's almost impossible to find people to review, finding native people that could *make* the glyphs is often completely out of the question

Metrics different scripts incompatible

- Already a problem with Vietnamese, certainly a problem with Arabic, Khmer, etc.

But:

- Solution exists in OpenType in form of BASE table, but almost unsupported now

Don't put it all in one font

- Pushing against limits of renderers
 - e.g. Kerning bug in Pango that was suddenly triggered by DejaVu
 - “Safer” to use smaller fonts
- Pushing against limits of editors
- Hard to manage such a big font
- Font files become huge
 - Would probably result in performance issues with development tools like svn diff

Don't put it all in one font (continued)

- Putting script X in DejaVu forces me to use those glyphs from DejaVu for that script if it's my default font for Latin
- At least split the fonts up for the different scripts...

Splitting the fonts?

- Inevitably we should split the fonts if new scripts keep being added, due to limits in the font format
- But a set of fonts is harder to maintain
 - Difficulty of keeping common glyphs and font data synchronized
- Users will have dozens of entries in their font menus for each script
 - Not really user friendly...
- Splitting fonts also not trivial due to common glyphs etc.

Combining the split up fonts?

- Combine the different scripts in meta-fonts (like Sans, Serif and Monospaced)
- Could be done with Fontconfig

Fontconfig a solution?

- Users would have to edit their fonts.conf files manually, or their distribution package management systems would have to do it for them
- As a font, we simply cannot rely on Fontconfig to render our fonts in a usable way
 - Linux programs bypassing Fontconfig, Windows, MacOS...

Fontconfig a solution?

- Different computers, different fontconfig configurations
 - Same document, different fonts used, often resulting in different layout

Conclusion

- ?