
Typesetting classic Jewish texts

Leonid Dubinsky

Table of Contents

TODO	2
Font Formats	3
TrueType, Type1 and history of OpenType	4
Font Rendering	5
Fonts	6
Glyph Layout with Vowels and Cantillation	7
People	8
Intermediate Format	9
PDF generation	10
XSL-FO engines	11
FOP	11
XEP [XEP]	11
xmlroff [xmlroff]	11
passivetex	11
Notes	12
Bibliography	12

TODO

- Make bibliods of the uri class active!
- FOP
- When do we need this
- Examples of the problem with images
- History of contacts with the SIL font guy

Font Formats

The problem: <http://www.tanach.us/Tanach.xml#Installation> [<http://www.tanach.us/Tanach.xml#Intallation>].

To show on screen (and print) acceptably texts with vowel points and cantillation (and without them), we need good fonts. It would be nice to be able to force the browser to show a page using a specified font, and there is even a declaration for that in CSS (TODO), but it is not universally supported by modern browsers/operating systems, so the user will need to install a decent font if he is not satisfied with the quality of the pre-installed ones. Font installation is easy, but we need to provide the fonts to install.

We need to be able to install fonts in cross-platform formats.

It happens that there are both a vowel and a cantillation sign under a letter (vowel is printed first in such cases (TODO)). How can we make sure that vowel and cantillation sign do not overlap?

Apple developed font format that allows inclusion (in form of tables) of instructions about combining glyphs. This format used to be called GX, and is now called AAT - Apple Advanced Typography [AAT]. It is currently supported only on MacOS (there are plans to support it in HarfBuzz [HarfBuzz] in the future (TODO)), so we can not use it.

SIL International is developing AAT replacement for Windows - Graphite [Graphite]. This format allows inclusion into the font of a program in a special language. This program then chooses and places glyphs. This format is supported on Windows and Linux, but requires installation of programs in addition' to fonts, so we can not use it.

TrueType, Type1 and history of OpenType

Microsoft and Apple developed OpenType format [OpenType]. This format has various tables and is more expressive than Type1 and TrueType [TrueType], but less expressive than AAT. Applications rely on a platform library to render text well for a complex writing system (like ours) - or do it themselves. On windows, Uniscribe [Uniscribe] is such a library. We need to find out, how well it really supports, as they said, "the use of kerning for adding space to bases with diacritics (Nikud or Teamin)". On Linux, this is done by Pango [Pango], and it, it seems, does not deal with cantillation at all - just with consonants and vowels [PangoH]; we need to contact them, find out what the situation is and possibly change it. Firefox (the only program that we care about) uses Pango, which works on Windows (via Uniscribe), and on MacOS.

Font Rendering

Firefox: Pango everywhere?

FreeDesktop [<http://www.freedesktop.org/>] Text Layout Working Group [<http://www.freedesktop.org/wiki/TextLayout>]

Fedora Fonts SIG [<http://fedoraproject.org/wiki/SIGs/Fonts>] QA [<http://fedoraproject.org/wiki/SIGs/Fonts/QA>]

HurfBuzz [HarfBuzz]

Cairo [<http://cairographics.org/>] graphics

TextLayout [<http://freedesktop.org/wiki/TextLayout>] 2006 [<http://live.gnome.org/Boston2006/TextLayout/>] 2007 [<http://www.freedesktop.org/wiki/TextLayout2007/>]

Open Font Library Wiki [http://openfontlibrary.org/wiki/Knowledge_Resources]

Fonts

- SIL Ezra
- Cardon?
- Gorkin?!

Glyph Layout with Vowels and Cantillation

TODO XXX's algorithm.

Gorkin's algorithm.

There is a book on fonts: Fonts & Encodings by Yannis Haralambous. [Hara]. Despite the raving review [http://www.oreillynet.com/xml/blog/2007/10/fonts_encodings_by_yannis_hara.html], I was underwhelmed by it. I wanted to find out if it is possible to encode the vowel/cantillation placement logic into an OpenType font - and did not find the answer in the book.

Which is not really surprising, since the book's author is also the author of Omega project (Unicode in TeX), about which a very informative text [http://www.valdyas.org/linguistics/printing_unicode.html] about printing Unicode in 2002 says:

- However, Omega is very much a failure. Its creators have been guided by Principles. They were conscious of the Desirability of Flexibility. They Knew about the Demands of Fine Typesetting of Complex Scripts. ... You cannot run to the manual, because the manual is a very interesting piece of academic prose about the difficulty of the task, but useless for a mere user.

I *did* find out about a useful tool for OpenType font jobs: TTX [TTX].

Does kerning work through vowels?

People

Intermediate Format

Repeatability and hand-finishing: contradictory requirements?

PDF generation

XSL-FO engines

For real typesetting of a tree of texts we'll need to generate our own PDF anyway. For typesetting papers and other project documentation, almost anything will work.

For typesetting of the Hebrew text with vowel points and cantillation, we have:

FOP

No support for OpenType fonts

XEP [XEP]

Ignores OpenType GPOS/GSUB table, so useless for typesetting Tanach. Attempts to contact support for clarifications failed.

xmlroff [xmlroff]

- Does not support regions other than main.
- Excellent with OpenType (uses Pango).
- Excellent support (see <http://xmlroff.org/ticket/131> for an example)
- There seems to be some issue with embedding the fonts and display on MacOS/Windows.

passivetex

antennahouse and other commercial::No breaks for non-profit

Notes

There are rumors that Pango processes cantillation correctly - possibly, with good fonts? We need to ascertain - with Behdad? - that we do not need special support from Pango, and that expressive power of OpenType is sufficient.

InDesign [InDesign] and its storage format INX [INX] are something to think about in the context of Outside-In XML Publishing [Outside-In].

"Typesetting Hebrew Cantillation".

Bibliography

[AAT] *Apple Advanced Typography (AAT)*. Apple. <http://developer.apple.com/fonts/TTRefMan/RM06/Chap6AATIntro.html>.

[Graphite] *Graphite*. SIL. http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=GraphiteFAQ.

[TrueType] *TrueType*. <http://en.wikipedia.org/wiki/TrueType>.

[OpenType] *OpenType*. Wikipedia. <http://en.wikipedia.org/wiki/OpenType>.

[Uniscribe] *Uniscribe*. Microsoft. <http://www.microsoft.com/typography/otfnt-dev/hebrewot/features.aspx>.

[Pango] *Pango*. <http://www.pango.org/>.

[PangoH] *Pango Hebrew*. <http://cvs.gnome.org/viewcvs/pango/modules/hebrew/hebrew-shaper.c?view=markup>.

[XSL-FO] *XSL-FO*. Wikipedia. <http://en.wikipedia.org/wiki/XSL-FO>.

[Anvil] *Anvil Toolkit*. Dave Pawson. <http://www.dpawson.co.uk/nodesets/entries/070709.html>.

[Prince] *Prince*. <http://www.princexml.com>.

[GoogleBooks] *Google Books PDF*. <http://www.imperialviolet.org/binary/google-books-pdf.pdf>.

[HarfBuzz] *HarfBuzz*. <http://www.freedesktop.org/wiki/Software/HarfBuzz>.

[Hara] *Fonts & Encodings*. Yannis Haralambous. <http://www.amazon.com/Fonts-Encoding-Yannis-Haralambous/dp/0596102429>.

[XEP] *XEP*. RenderX. <http://www.renderx.com/RenderX>.

[xmlroff] *xmlroff*. <http://xmlroff.org/>.

[InDesign] *Adobe InDesign*. http://en.wikipedia.org/wiki/Adobe_InDesign.

[INX] *INX*. http://avondale.typepad.com/indesignupdate/2005/08/what_the_heck_i.html.

[Outside-In] *Outside-In XML publishing*. <http://2007.xmlconference.org/public/schedule/detail/249>.

[TTX] *TTX*. <http://www.lettererror.com/code/ttx/index.html>.

