RespExT the TEXniX

John Haltiwanger
mono here

A brief proposal for an interdisciplinary software study

For a thesis, I would like to propose a critical and operational engagement with the TEX typesetting engine.[1] What issues of materiality apply? What aspects of TEX, if any, are unique?

Intertwined are threads of progress (through the macro package ConTEXt) and collaboration (through the operational component, which is primarily concerned with establishing the infrastructure for a "pure peer" journal).

Critical Component

TEX has existed as an electronic typesetting engine for a three decades. As an ecosystem, TEX is defined by its self-documentation. The source itself is open, but beyond that the source of individual documents are frequently provided. This allows a culture of technique diffusion and could qualify as a virtuous process as per Benkler. It also means the TEX project could be considered a self-documenting electronic typesetting assemblage, which to my mind provides a unique opportunity for investigating the materiality of electronic type. Whereas an experienced typographer can determine the processes (or at least parameters) used to create output on a page, this capacity requires years of training. On the other hand, after mastering TEX to a certain degree, it is possible to look at source documents and learn the exact parameters utilized to generate various outputs.

TEX documents exist in multiple stages of materiality:

- *As a source document of marked up text.*
- *As intermediately processed files that are generated during document compilation. (During this stage, any errors in the source document result in an interactive compiler prompt.)*
- *As a processed output file (DVI, PS, PDF)*

[1] *Though the majority of investigation will deal with ConTEXt rather than TEX, the project will hope to always ground the former in reference to the latter.*

- *As ink on a printed page.*
- *As [potentially interactive] pixels on a screen.*

(The operational engagement will incorporate an additional materiality: as a pre-TEX source format, most likely reStructuredText. This functionality is provided by a translation layer, most likely pandoc, adding another degree of intrigue to TEX's materiality. Many other formats besides TEX can be output through pandoc, giving a degree of format parity rarely held by TEX documents.)

Though TEX is powerful, it has also continued to evolve. Through macro packages such as LATEX and ConTEXt, TEX has become considerably easier to use. Furthermore, developments such as XeTeX and LuaTeX are pushing the envelope in terms of international support (Unicode, non-Western text formatting, etc). As ConTEXt will be utilized in the operational component of the project, combined with the fact that it is under the most heavy and promising development at the moment, I expect that ConTEXt and LuaTeX will be central to the critical engagement (perhaps even to the extent that it becomes a software study of ConTEXt more than TEX). The capacity of ConTEXt to generate electronic documents, for instance, makes for another layer of materiality: `hypertext.' To put it simply, ConTEXt manuals often incorporate not only page-level links to the index, table of contents, and chapter-level section lists, but to a search function as well. This `hyper' level functionality has real implications not only for electronic typesetting in general, it also decreases traditional obstacles that define `learning curves' in the actual absorption of the ConTEXt (sub-)assemblage.

By engaging actively with it's most active macro package (ConTEXt) issues of revision and editability can be investigated on the material levels of that define TEX documents. By engaging with evolving software, the research accepts and acknowledges that its specific critical results are somewhat tied to the version(s) of ConTEXt that are encountered during the study's unfolding. However, a theoretical framework for understanding the characteristics and potentials unique to TEX, today, will remain useful in spite of the inevitable progress of ConTEXt.

Preliminary Reference List

- *Chun, Wendy Hui Kyong. (2008). Control & Freedom.*

  Materialist framework for investigating the `light' of electronic typesetting. Also for its investigation of gender and race in relation to technology.

- *Gitelman, Lisa. (2006). Always Already New. For its historically rooted analytical framework of the archive.*

- *Hagen, Hans. (2009). ConTEXt Mk.IV Manual A sequential documentation of all the changes made to ConTEXt through the evolution of LuaTeX. Excellent resource for discussing changes to TEX/ConTEXt.*

- *Hayles, N. Katherine. (2002). "Text is flat, code is deep". Another avenue for materialist analysis.*

- *Hoekwater, Taco, et al. (2009). LuaTeX Reference Manual. The constantly-evolving reference to LuaTeX provides insight into the maturation of this new engine.*

- *Kinross, Robin. (2004). Modern Typography By engaging typography through the lens of a "critical history," this text should provide important information for clarifying the uniqueness (or non-uniqueness) of electronic type.*

- *Knuth, Donald. (1984). TEX: The Program. For its detailed description of the TEX typesetting engine.*

- *Plaice, John, et al. (2008). "Multidimensional text". Electronic text as a tuple; interesting framework.*

Operational Component

This part of the project developed through a thought experiment based on Geert Lovink's proposal to found a "post-peer" review journal during the last class of New Media Practices 2009. My main concerns in pursuing a theoretical degree in new media involve issues of collaborative processes, and so I naturally applied some of the technical processes I've been contemplating to solving the problem of a "post-peer" journal. The first order of business will be establishing the concept of "pure peer," which emphasizes that what is proposed is not so much "post-peer" but a new form of peer review that more closely follows a P2P model.

This operational component will be concerned with developing a web application that allows near-universal export formats for its documents. One means of accomplishing this

is by using an intermediate source language, something that is not TEX and not HTML/XML/etc. (Most likely pandoc will be used as a document translator, and reStructuredText will be used as the base language). A major reason for this plethora of export options is that ConTEXt is included as an output format, opening avenues for beautifully typeset PDFs that can target not only paper, but the screen as well. By publishing such documents, it is hoped that traditional barriers to online journal (not the least of which is the general awful-ness of text in the browser) will be mitigated by professional, and in some ways "next level," presentation.

Issues of collaboration are raised as we contemplate the potential for multiple authorship at the journal. What modes of collaboration should be available and/or encouraged? By utilizing a plain text format such as reStructuredText, we allow for atomic contribution tracking by tying in `git` *version control. Whether collaboration becomes commonplace or not, the revision history will always be available to provide another layer for investigation. Adding version control "materializes" the medium of the source code in a useful way.*

The project involves the following components:

- *ConTEXt: for its modern features and flexibility*
- *pandoc: for it's ability to transform documents*
- *Waves: (that's Ruby Waves) for it's robustness and modularity as a web application platform*
- *RDF/A: because (TEX) PDFs still contain no metadata, RDF/A will be used along with md5sums to convey important metainfo about documents published using the platform.*
- *git: because it allows for atomic version control, and also provides elements that can be considered virtuous (see 'Git Virtue?' on the MoM blog for more). basically this mechanism can show who contributed what to which article.*