

Math in ConT_EXt roadmap

1. After playing with math support for more than a year, we have come to the conclusion that it is time to move on. We have already discarded italic correction and now are replacing rules with extensibles. Much was already in place (and applied) but experiences with type one antykwaz made us review some OpenType fonts. Not using rules makes some of them look better. The effect is subtle and probably not ams compliant, but we think that it will work out well for simple math like fractions of decimal numbers. Consequently, we have added to our shrinking to-do list the burden to investigate whether we can remove those obsolete code paths from the engine. After all, who needs italic correction, who prefers ugly rules to beautiful glyphs, and who understands all these font parameters? Furthermore, after all these years, we don't expect OpenType font and Unicode math technologies to improve much; we don't know if Microsoft is developing their technology further at all. Therefore, we are confident that what we are doing is the way it should have been done when math was upgraded. Hopefully users will notice the improvements.
2. Math also means physics and units (that topic was brought up recently on the list by Gavin). Therefore, because we're in cleanup mode, we decided to eliminate some more. With iso now in place for a long time, we are going to ignore the existence of the inch as unit from now on. The unit will probably remain in the engine for nostalgic reasons, but it will no be accepted in MWE. Instead, we will provide some more modern, culturally correct, kid-friendly units that we will use in examples, manuals and such. Because the four-person strong team dealing with this wants to avoid making mistakes, we will go through a careful and scientifically sound process of calibration first, using a selected tex savvy audience. We expect these new units to be stable a month from now. Believe it or not, in the process of documenting all this, we found a buglet in the new math dimension spacing, so it has already paid off. Expect to hear more in a month or so, and enjoy your inches as long as you still can. In case you wonder how this relates to math other than mentioned: the math subsystem has 'mu' as adaptive unit, and that inspired is to come up with one for text (in addition to two new more or less fixed units).
3. The math family model is a fundamental concept in T_EX but we think we can do without. First of all, OpenType math fonts have (design) script and scriptscript sizes built in, so for that we have one family. Second, only full bold (heavy) makes sense as companion for regular math which is something that in practice we can support otherwise. So, this makes us consider dropping families altogether which then provides (mem) space for even more classes or dictionaries. If we nevertheless decide to keep families, we can certainly go with less than we have now, maybe two (or four if we want to be generous and also resemble original tex) of them is enough. We cannot imagine users wanting more. As a side note: completely divorcing families could make the math engine a bit leaner. It is hard to explain and users only care about the outcome. So more on this later.
4. Another path to explore is to identify the few building blocks that are needed for typesetting math, and then doing a bit more at the tex end. Of course that would nil quite some earlier effort, which is a bit frustrating, but still . . . maybe the math engine can be reduced to a fraction of what is is now.
5. When we look at the math fonts and some characters in there, we sometimes wonder what makes sense. For some, searching in e.g. arXiv brings no hit. Basically we have obsolete math symbols and currently used one. That made us think about ancient math versus modern math, just like there is ancient greek and modern greek. Because math is a script one can wonder about obsolete math dialects with symbols just like there are plenty deal scripts in Unicode. We already are working on dictionaries but another axis is useability.
6. We no longer have the small / large extensible family model so we can simplify delimiters in the engine. Not something users should worry about.
7. We're not sure why math is considered stable because everything moved forward. Therefore we're preparing a bid for extra math symbols as needed in modern explorative and daring math thesis. When symbols are really used, and we have proof of that, it should be possible to get them un Unicode, just like all these emoji. We welcome input and as an example of currently faked symbols we added some to the distribution as easter eggs. One example:

Mikael got contacted by a stressed student working on a thesis on probability. This student needed to typeset the characteristic function of a random variable X with density function f_X , and it was insisted to use another notation than the (wide) hat, that was already used for something else. For this reason the `\widerandomhat` was introduced,

$$E[e^{itX}] = \widehat{f_X}(t), \quad E[e^{it(X_1+X_2)}] = \widehat{f_{X_1} * f_{X_2}}(t).$$

Naturally, it is automatically scaled, just like the ordinary wide hat

$$a + b + \widehat{c + d} + e + f \neq \widehat{a + b + c + d + e + f}$$

Once the thesis is printed, we will contact the Unicode Consortium to suggest that it gets a slot.

8. Our most ambitious project is a reverse engineering one, which is why it is conducted at the engineering faculty of the Dnubn university (we cannot reveal the real name yet). In math articles one can find visualizing like $x \leftarrow x$ and there are plenty of \TeX commands that have arrow or hook in their names. If you look at the names of math symbols plenty are kind of weird. We think it is not natural and are considering a ‘natural language math input’ project, where you tell what it is and get the symbols you expect. For that we need to analyze typeset math and from the context as well as visualization derive a dataset that we can feed into a machine learning subsystem that then can be used to turn input into type. We have several stages in mind spanning years but it can be fun. Think of it like ‘untagged math’ which then of course results in ‘untagged pdf’, but better!

Mikael & Hans