New dimensions: Edith and Tove

Willi Egger, Hans Hagen, Edith Sundqvist, Mikael P. Sundqvist

Male dominance

When you start using TEX you can't get around the fact that it uses dimensions. You have to set up a paper size, configure a line width, tell it what font size to use, etc. As with many techniques that evolved in different countries the way to express a dimension can be done differently. In Europe we like to talk in centimeters (`cm`) or millimeters (`mm`) and in the United States it's all about inches (`in`). Typographers all over the world speak in terms of points (`pt`), didots (`dd`), ciceros (`cc`) and picas (`pc`) while those messing around with digital typography prefer "big" (PostScript) points (`bp`). TEXies sometimes like scaled points (`sp`) as 1 `sp` is the smallest internal representation of a unit. When someone talks "points" you can't be sure if it is big points or TEX points because the `pt` unit is often used for both.

There are also font-related units, like the popular em width (`em`) and ex height (`ex`) and there is even a pixel unit (`px`) that can be set to some resolution but that one is rarely used. There is also a math unit (`mu`) that scales with the math font in use.

All units are internally scaled points and one real point is 65536 scaled points. That means that when a unit is entered it gets mapped onto this internal scaled point quantity.[1]

For a while we had the new didot and new cicero but in LuaMetaTEX these were dropped because no one used them. On the contrary, the recently introduced (Don) Knuth unit (`dk`) is quite convenient and we use it as a convenient offset for a so-called TEX page environment, which we use a lot in testing math functionality (Hans Hagen, A new unit for LMTX: The `dk`; `tug.org/TUGboat/tb42-3/tb132hagen-dk.pdf`).

From this summary we can observe that there are three units that are names: Didot, Cicero and Knuth. But do you realize that these are all males? That can't be right and should be fixed. If you look at user styles (or questions on support platforms) you will also notice that in spite of standardization, the inch (`in`) has not been replaced by its more correct metric counterparts. Okay, that might be due to the fact that there is no meter as unit but using smaller dimensions (`mm` and `cm`) makes more sense, also for internal accuracy reasons. That said, it is about time that we eradicate the inch or at least come up with something more metric.

So there you have it: we need some female units that correctly stay within the metric domain! In order to convince users to drop the inch the first new unit somewhat relates to it: one Edith (`es`) is the median of the widths of thumbs of BachoTEX 2023 attendees. One can argue that this is somewhat arbitrary and indeed it is. In order to get a decent value we use a discrete measurement device that groups thumbs into 15, 20, 25 and 30 mm intervals. A 10mm interval is unlikely to get many hits unless the TEX ecosystem suddenly became very easy to use and toddlers get interested in it as a game.

Rule of thumb

If we talk in terms of one Edith, we should keep in mind that at any point we can decide to re-calibrate that unit. If we end up below 25mm we probably have quite some young and/or old users in the sample set. So, in order to have a constant value, the community has to make sure that TEX (and preferably ConTEXt) usage is nicely distributed. Now, of course at BachoTEX we are quite tolerant, because also Plain and LATEX users are sampled. Also, given that this sample of the TEX community is skewed to older users, one can wonder how that influences the initial value. It is up to the ConTEXt group to decide when and where to re-calibrate at a later moment. After all, we have to keep the narrative that ConTEXt is unstable and evolves alive, and occasionally updating a unit fits into that narrative. If you think that this kind of research is somewhat flaky, keep in mind that probably all research related to typography is kind of subjective and somewhat unreal. And BachoTEX being tagged as 'conference' adds a lot of credibility.

The Edith (▬) makes a nice unit for margins, but it is a bit large for offsets, so we also need a female counterpart for the Knuth (▬). This is why, just like a centimeter (▬) has a smaller companion in millimeter (▮), the Edith has a companion Tove (▬). In terms of points one Tove is 7.11317pt, while a Knuth is 6.43985pt. It is surely just a coincidence that the value of one Tove in points is about the age of Tove when she became aware that her dad was a ConTEXt fan. In terms of points one Edith is 71.13177pt which, ignoring the unit, comes close to the average age of those who have attended BachoTEX more than 10 times.

Originally presented at BachoTEX 2023, and to be published in the corresponding *GUST Biuletyn*.

[1] If you go back to the early days, there are even cases where you want to talk in terms of true units. Those are not affected by original TEX's magnification factor (`\mag`) but in LuaMetaTEX we dropped that factor and therefore also these true units became obsolete.

doi.org/10.47397/tb/44-2/tb137egger-edith

The implementation of these units in LuaMetaTEX is not that hard, simply because scanning for these dimensions happens in few places: when scanning dimensions, and in a Lua helper that converts a string to scaled points. At the ConTEXt meeting where we implemented the Knuth, there was some trial and error involved in order to get the right numerator and denominator. One `dk` is 422042 scaled points which brings us to a numerator 49838 and denominator 7739. Except for scaled points, the fraction gets multiplied by 65536 and the amount. Most units have numerators and denominators with weird values, although 7227 jumps out.

unit	visualized	name	num	den
bp		big point	7227	7200
cc		cicero	14856	1157
cm		centimeter	7227	254
dd		didot	1238	1157
dk		knuth	49838	7739
es		edith	9176	129
in		inch	7227	100
mm		millimeter	7227	2540
pc		pica	12	1
pt		point	1	1
sp		scaled point[a]	1	1
ts		tove	4588	645

[a] This one is not multiplied by 65536, and has been greatly enlarged to be visible here.

When you consider these numbers it is good to realize that internally the engine uses a 32 bit number, split into two halves. There is a maximum, 16383.99998pt, so that (intermediate) calculations don't overflow. The last digit of what TEX reports when it computes a dimension as points is to be taken with a grain of salt. Here is how the Edith and Tove compare to their metric counterparts:

2.5cm	4661699	71.13188pt
2.5mm	466169	7.11317pt
1es	4661692	71.13177pt
1ts	466169	7.11317pt

In case you wonder if checking for yet another unit has drawbacks in terms of performance, we can guarantee LMTX users that they won't notice a performance hit. Even with these additional units the engine quite likely beats its predecessors in scanning units. And the impact on the code base is less than 20 short lines of trivial code so that goes unnoticed as well.

Calibration

In order to conduct the calibration we need a reliable measurement device and here we got lucky. The ConTEXt community has some unique craftsmanship amongst its members and Willi Egger made us a robust sampling device that can compete with those used by the ones that the International Organization for Standards uses: the Edithorial.

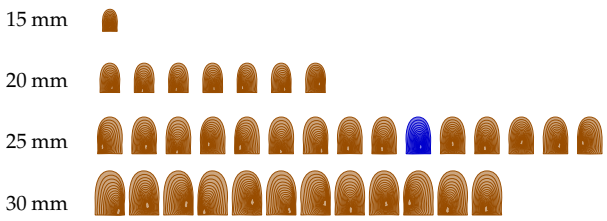


Figure 1: Results from thumb measurements at BachoTEX, with the median thumb marked in blue (grayscaled for print).

In addition to that, the ConTEXt Math Society, indeed the same one that brings you all these nice new math capabilities in LuaMetaTEX, provided the necessary statistical and mathematical underpinning to make the Edith and Tove believable units. So here are some more details.

We have found out that the Tove unit, 2.5 millimeters, corresponds to 7.1131744384765625 points. Let us find a decent rational approximation of this, with a small denominator. We do this by calculating the continued fraction, and we try a few steps to get something that is good enough.

We start by noting that the integer part is 7. We then use a calculator (in our case Wolfram Alpha) to compute

$$\frac{1}{7.1131744384765625 - 7} = \frac{1}{0.1131744384765625} \approx 8.835917486854523392207091816098.$$

This means that we get as a first possible choice

$$7 + \frac{1}{8} = \frac{57}{8} = 7.125\ .$$

We continue, and next note that

$$\frac{1}{8.835917486854523392207091816098 - 8} \approx 1.196290322580645161290322580645.$$

Thus, our next candidate is

$$7 + \frac{1}{8 + \frac{1}{1}} = \frac{64}{9} = 7.\overline{1}.$$

Here, the bar over the 1 indicates that 1 is repeating.

In the next step we calculate

$$\frac{1}{1.196290322580645161290322580645 - 1} \approx 5.094494658997534921939194741167.$$

The next candidate becomes

$$7 + \frac{1}{8 + \frac{1}{1 + \frac{1}{5}}} = \frac{377}{53} = 7.\overline{1132075471698}.$$

We continue, to get

$$\frac{1}{5.094494658997534921939194741167 - 5} \approx 10.582608695652173913043478260870.$$

The next approximant becomes

$$7 + \frac{1}{8 + \frac{1}{1+\frac{1}{5+\frac{1}{10}}}} = \frac{3834}{539}$$

$$= 7.\overline{113172541743970315398886827458256029684601}.$$

For the next step we have

$$\frac{1}{10.582608695652173913043478260870 - 10} \approx 1.716417910447761194029850746269$$

so the next approximant becomes

$$7 + \frac{1}{8 + \frac{1}{1+\frac{1}{5+\frac{1}{10+\frac{1}{1}}}}} = \frac{4211}{592} = 7.1131\overline{756}.$$

Since this one has such a nice short repeating set of decimals, we fell for it, and quit here. The next approximants would be

$$\frac{12256}{1723}, \frac{20301}{2854}, \frac{28346}{3985}, \frac{48647}{6839}, \frac{466169}{65536},$$

where the last one exactly equals what we started with, 7.1131744384765625. Before we continue, we mention that $[7; 8, 1, 5, 10, 1, 1, 2, 1, 1, 9]$ is a more compact way to write the continued fraction above.

One could perhaps first think that multiplying the rational number by 10 would yield a very similar continued fraction, but that is not the case. In fact, the continued fraction for 71.131744384765625 is given by $[71; 7, 1, 1, 2, 3, 1, 3, 1, 3, 1, 3, 2]$. This put us in a bit of an awkward situation. Do we want a nice approximation for the true value, or do we prefer to have `es` to be exactly 10 times as large as `ts`? If we go for the latter, we could take 42110/592. We calculated the approximants though, and got

$$\frac{489}{7}, \frac{569}{8}, \frac{1067}{15}, \frac{2703}{38}, \frac{9176}{129}, \frac{11879}{167}, \frac{44813}{630}, \frac{56692}{692}, \frac{214889}{3021}, \frac{271581}{3818}, \frac{1029532}{14475}, \frac{2330845}{32768}.$$

When we saw this, it was irresistible to define `es` as

$$\frac{9176}{129} = 71.\overline{131782945736434108527}$$

and then to define `ts` as

$$\frac{9176}{1290} = \frac{4588}{645} = 7.1\overline{131782945736434108527}.$$

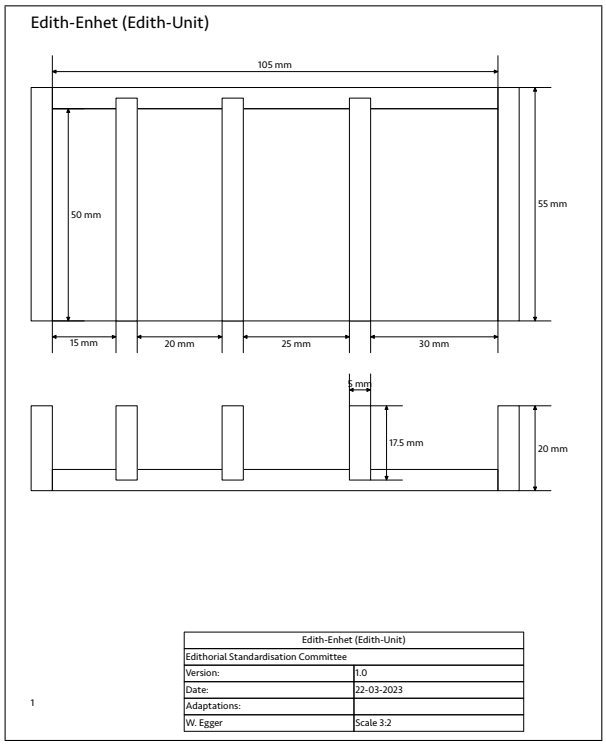


Figure 2: A TEX-community-worthy edithorial for measuring the Edith.

The edithorial device

The design of the edithorial also involved some research. Of course there was some discussion about the right way to sample thumbs and those who have attended BachoTEX and ConTEXt meetings will not be surprised that Willi is responsible for this. He presented us with a drawing (figure 2) that we immediately agreed upon.

Willi then sat down and made a prototype (figure 3) in order to see if sampling would work out. Knowing that the device would be stored under harsh conditions in the university city of Lund in Sweden, it had to be sturdy Polish oak and after being brought to precision it underwent first an iron acetate treatment and after that a furniture oil (tung oil) treatment as can be seen in figure 4. Even with TEX being digital we cannot get around physical devices for measuring digits. And with TEX operating in nanometers we have to fit in.

Some double checking

There is one question we have to answer before we dare to use the Edith (`es`) and Tove (`ts`) as offsets next to a Knuth (`dk`) and that is: in what box does Don's thumb fit? After all, we need to assign some more weight to his thumb. On the Internet you can find images of Don Knuth sitting behind an organ

Figure 3: The prototype of the edithorial.

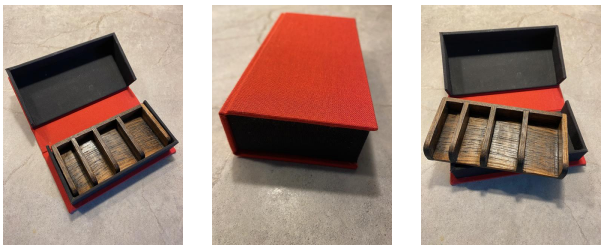


Figure 4: The reference edithorial with protective cover.

but for reasons of copyright we cannot show these, but one thing we can be sure of is that his thumb is not wider than a key of that instrument, because according to the Wikipedia page `Musical_keyboard`:[2]

> Over the last three hundred years, the octave span distance found on historical keyboard instruments (organs, virginals, clavichords, harpsichords, and pianos) has ranged from as little as 125 mm (4.9 in) to as much as 170 mm (6.7 in). Modern piano keyboards ordinarily have an octave span of 164–165 mm (6.5–6.5 in), resulting in the width of black keys averaging 13.7 mm (0.54 in) and white keys about 23.5 mm (0.93 in) at the base, disregarding space between keys.

[2] Notice how metric measures win over inches here!

Figure 5: The 2019 lecture: Pi and *The Art of Computer Programming*.

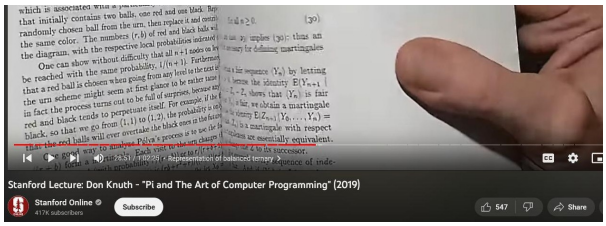


Figure 6: The 2014 lecture: (3/2)-ary Trees

This definitely keeps Don's thumb out of the 30mm bucket. When we zoom into these images it seems also unlikely that the thumb will go to the 20mm bucket, but in the end the only one who can answer this is Don Knuth himself. And because he's behind an email firewall we don't dare to ask him. So more research was needed and after a brainstorm session we decided to rely on a public visual that any TEX user should be familiar with: the yearly Christmas lectures (figures 5, 6).

And because we know which books the thumb is on, we can calculate the bucket by comparing the dimensions: on one case we use the paper size as a reference, on the other case we use the interline spacing of the book as reference!

In Figure 7 we show a close-up of the thumb and the page. We have divided the image into a 100×100 grid, but the aspect ratio of the image is $3 : 2$, so we need to compensate for that. We estimate that the interline space of the text is 8 grid lines high, while the diagonal line measuring the width of the thumb is 12 grid lines wide and 42 grid lines high. This means that the thumb-interline space quotient is given by

$$\frac{\sqrt{(12 \cdot 3/2)^2 + 42^2}}{8} \approx 5.71.$$

Next, we need to know what interline space is used. We should probably know this by heart, but as we do not, we instead downloaded one of the prefascicles of *TAOCP* volume 4. We cut out a square with sides of 5 cm, and added again a 100×100 grid.

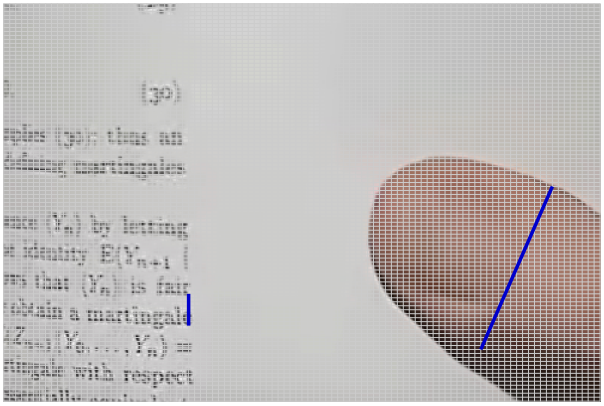
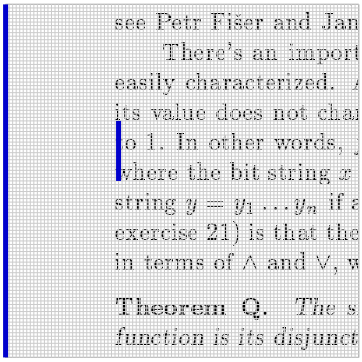


Figure 7: Close-up of Don Knuth's thumb.



We measured the height of two lines and got in return 17 grid lines. This means that the interline space is given by

$$5 \times \frac{17}{2 \cdot 100}\,\text{centimeter} = 0.425\,\text{centimeter}\ .$$

As a result we estimate that Don Knuth's thumb has the size

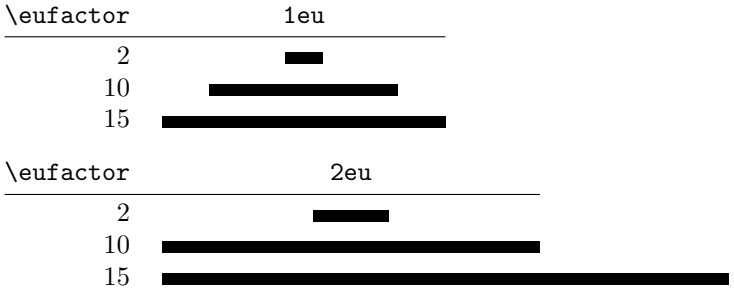
$$5.71 \times 0.425\,\text{centimeter} \approx 2.43\,\text{centimeter}\ .$$

If we're right about all this then the Edith will not be influenced by the grand wizard's thumb, so the well-calibrated (derived) Tove cannot be discarded for offsets as being less accurate (and stable) as the Knuth.

A modern relative unit

Since TEX showed up, a lot has changed when it comes to computers: the computers considered powerful in the early days now fit in your pocket. One disadvantage of these portable devices is that they have a variety of display sizes. A document can easily be generated again, adapting the layout to all these devices is a bit of a pain.

This is why we introduce a new dynamic unit, the `eu` or the European Unit, but one that can be changed by setting an internal register, `\eufactor`. Because that defaults to 10, one `eu` starts out as one `es`. A nice coincidence is that one can also read it as Edith's Unit.



We can set the factor in Tove steps between 1 and 50 so that we retain a reasonable accuracy. So, this relative unit stresses the sisterhood of these two new units because `1eu` is `10ts` and `1es`. This unit might also come in handy when writing manuals so you can bet that we will use it.

These units are modern in another way too. The popular game MineCraft has its own unit, a block, as (for instance) discussed on `minecraft.fandom.com/wiki/Tutorials/Units_of_measure`. For those using inches, one inch is 0.0254 blocks, so one block makes 39.3700787in. For those using metric system one cm equals 0.01 blocks or 0.16 pixels and therefore one block makes 40cm. These 40cm are 16 Ediths which means that the Edith is also a good introduction in the hexadecimal numbering system. Unfortunately LEGO bricks are defined in inches so there the inchers still have the edge. But Edith and Tove have an advantage in MineCraft, which is confirmed by observation. Just like some of TEX's units are actually defined using the inch paradigm, we could add units like `mb` for MineCraft Block being 16 Ediths. After all, implementing extra units is trivial in LuaMetaTEX. Let us know what you think.

How about MetaPost?

We not only have to deal with TEX but also with MetaPost, so from now on Metafun will also provide these units, which we can then use to properly draw thumbs as in figure 8.

While checking other units in Metafun we were reminded that they are there given as floats and not as fractions. We were amused to see

```
mm :=  2.83464 ;
cm := 28.34645 ;
```

which means that a `mm` is not exactly one tenth of a `cm`, and also that the rounding has been done by the even/odd rounding off rule. We decided to define

```
es := 71.13174 ;
ts :=  7.11317 ;
```

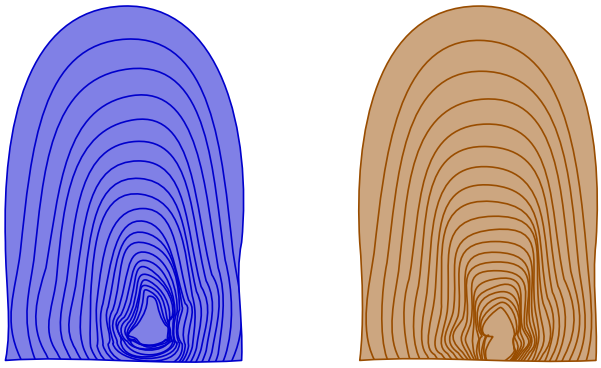


Figure 8: One can sign documents with these calibrated thumbs.

Wrapping up

In this article we discussed a few additional units that have been added to LuaMetaTEX. We've carefully chosen some names that not only compensate the male dominance in unit names, but also have a modern and fresh ring. The units are of course metric. The Edith (`es`) replaces the deprecated inch (`in`) and the Tove (`ts`) can be used for offsets as alternative to the Knuth (`dk`) that of course we will keep using alongside. The units are calibrated using an edithorial of which there exists a unique reference measurement piece. The standard has been established at the 2023 BachoTEX meeting and might be recalibrated at a future ConTEXt meeting when a new generation of users thinks that is needed.

Appendix: Overflow

When you enter a dimension in TEX and it is larger than 16383.99998pt or 1073741823 scaled points, an error message is shown and when you ask for help, that contains the sentence "I can't work with sizes bigger than about 19 feet". There is no `ft` unit in TEX, so the user has to do some conversion, maybe taking ones own foot into account.

Just like we had to adapt the error message issued when an unknown unit is used, we decided make the overflow message a bit more detailed. For that, we introduced the Theodore, where that unit is to the Edith what the foot is to the inch, and with one Theodore being five Edith. We now report this:

> I can't work with sizes bigger than about 19 feet (45 Theodores as of 2023), 575 centimeters, 2300 Toves, 230 Ediths or 16383 points.

So how did we come to this one? At the BachoTEX meeting the 18-month-old, always good-humored, Theodore was running around in the conference room and his little feet were carefully measured by his father Arthur Rosendahl (the self-appointed High Commissioner of Hyphenation and upcoming TUG president). Because the 19 feet are also an approximation, we rounded the Theodore to five Ediths. In addition we mention a few more maxima, so that the user gets a better impression how large TEX can go.

Mojca Miklavec, who gets her feet dirty by managing the binary build farm on the context garden, proposed a `th` unit but as there is no `ft` we didn't come to a conclusion yet. Although that unit would make a good default for text width, just like an `es` makes perfect left margin, and a `ts` a nice offset around framed content ...

◇ Willi Egger
BOEDE

◇ Hans Hagen
Pragma ADE

◇ Edith Sundqvist
Palettskolan
Division of Miró

◇ Mikael P. Sundqvist
Department of Mathematics
Lund University