

Kent Karlsson

Comments on draft UTR 25 (divided into two main sections: data file and report).

Data file:

Check all character names (in comments)! Examples of errors: for ranges the second character name is rarely complete; there are no DAGGER RELATION or DOUBLE DAGGER RELATION character names.

Some ranges are oddly divided, like A..K, L, M..Z.

Only 1, 2, 3 superscript digits are included; all the super- and subscript digits should be included, as should superscript n and superscript i (and reasonably also superscript j) and super/subscript parentheses, +, - and =, which, IIRC, were included for compatibility "math" use in terminal emulators. Similarly only some VULGAR FRACTIONS are included, while all of them should be included, and also the FRACTION SLASH.

Only some currency signs are included; all should be included on an equal footing (or none).

CAPITAL LETTER O WITH STROKE isn't included in the data file. This letter is the original (and still valid) denotation for the empty set, and should be covered.

There is a reference (within the data file) from OHM SIGN to LATIN CAPITAL LETTER A WITH RING ABOVE, but the latter character is not included in the data file. Indeed, the "angst" [sic!!] "ISOTECH" character should be mapped to LATIN CAPITAL LETTER A WITH RING ABOVE.

MICRO SIGN should be deprecated for math use, if at all included in the data file, and the ISONUM "micro" should be mapped to GREEK SMALL LETTER MU.

Should COMBINING LOW LINE and COMBINING OVERLINE really be included?

LATIN SMALL LETTER DOTLESS I should not be included; it is not for math use other than in words of Turkic origin used in math identifiers.

It is not related to TeX's `\imath` (and `\jmath`; a.k.a. `\i` and `\j`) which are just dotless glyphs for putting accents on top of i and j. The glyphs in TeX are by default italic, as is the default for all letters in math mode; but the font for these can be changed, to get upright, bold etc. glyphs. The "Soft\_Dotted" property (see separate proposal) should be used instead (unless you want to allocate a host of `<font>ed` dotless i:s and j:s).

There are a number of mappings for ISO\* math characters. Since TeX (with some extra packages) is very popular for typesetting math,

mappings for TeX/mathmode (incl. popular packages) should also be provided.

Some 5-digit character numbers (for CIRCLED characters) have a leading 0, which they should not have.

Report (page numbers are to the PDF as posted):

TOC (in PDF as posted): level 3 headings should be indented a bit more...  
There should be no text on conformance in an informational document.

page 6:

Re. letters, U+00D8 (LATIN CAPITAL LETTER O WITH STROKE), upright or italic, can legitimately be used to denote the empty set.

Re. combining diacritics, note that i:s and j:s are (should be, see separate proposal and below) Soft\_Dotted. This should be noted in this text. The "time derivative" on an i or j variable is either a bad idea, or would need two combining dot above (one to make them hard-dotted, another to denote the derivative).

(middle paragraph) The word MATHEMATICAL is missing (twice) in the last line.

"strokes of a character" -> "strokes of a base glyph" (characters don't have strokes)

"double struck" -> "double-struck"

"considerably" -> "considerable"

"sometimes", "less often"; but which is the "most often" used form??

page 7:

"their the"??

VS1 is not used to select negation form.

"2287" -> "2278"

page 8:

"its", "their"; unclear which case is really referred to.

Both solidus overlay and vertical line overlay can be used to indicate negation.

page 9:

(middle of page) Variation selectors are not used to select between negation stroke variants.

page 10:

(top of page) The "taking of limit expressions" is not a Unicode property, but is done via a higher level protocol.

(bottom of page) Please forget (and don't use) the terms "bra" and "ket". They are silly and meaningless. No characters are named that way.

page 11:

The white square brackets are now disunified between CJK and math. It's a pity that disunification did not extend to tortoise shell and lenticular brackets.

page 15:

table: in the middle line the straight and slanted entries should be swapped.

page 16:

mglyph???? what?? how??

page 17:

table appears to be incomplete

page 18:

table appears to be incomplete

page 19:

table appears to be incomplete

page 20:

clause 3.2: letters are commonly used as relation symbols in logic, especially for relation variables, written infix in expressions. E.g.  $xRy$  or  $x R y$ . Usually written in italics, since they are usually variables. But this notation can also be used for defined relations. I.e. relations need not be represented by "class R" symbols, letters (in various styles) do just fine! (And it needn't be single-letter names either.)

"imaginary part" -> "imaginary unit"

Mention that CAPITAL O WITH STROKE can be used to denote the empty set.

"Standard" math function names need not be seriffed, sans serif

works just fine.

page 21:

the quoted notation for open interval is perfectly logical, much much much more so than the mixed square and round parens convention!

page 22:

"automatic ligature formation": doing that at the character level is very ill advised, and should not be done! Nor should it be referenced in a document on math support!

The "math keyboards" paragraph goes overboard re. keyboard modifier keys functionality! The paragraph should be deleted.

IMEs aren't mentioned, but could be used for typing math characters.

No particular software company should be mentioned in the UTR! (Nor the details in that paragraph, which leave much to be desired.)

"plain text math": this idea should be deleted completely from the document! It hints about something that does not exist.

Clause 4.3 also goes out on a speculative limb, which is quite unnecessary in a UTR.

page 24:

See comment on TOC above.

----\*\*\*----