Comments on the Public Review Issue #169 "Glyph Variation of Double Oblique Hyphen"

Karl Pentzlin − karl-pentzlin@acssoft.de − 2010-07-29

In L2/10-162 »Proposal to encode a punctuation mark "Double Hyphen"«, a new character "U+2E3C DOUBLE HYPHEN" was proposed, the representative glyph of it being a horizontal form like a short equals sign.

This character was not accepted by UTC #123. Instead, the Public Review Issue (PRI) #169 was issued: *"Recently, the UTC was presented with evidence that indicates that the DOUBLE OBLIQUE HYPHEN is used in both oblique and horizontal forms. Therefore the committee is considering adding an annotation to U+2E17 DOUBLE OBLIQUE HYPHEN, indicating that it may appear in either an oblique or horizontal form. Public input on the suitability of this is being sought."*

Doing so would have several disadvantages.

1. The PRI text cited above implies the question: "Do the users of the U+2E17 DOUBLE OBLIQUE HYPHEN have a problem if their character is displayed by a horizontal glyph?"

    a. The Unicode standard says (on p.64 on the 5.2 edition):
        "D2 – *Character identity*: The identity of a character is established by its character name and representative glyph in the code charts."
        Thus, any user of the character may expect to get an oblique glyph when using this character, even if they would tolerate a horizontal one.

    b. In consequence of this, a lot of current widespread fonts employ an oblique glyph:

        Arial x⸗x
        Cardo x⸗x
        Code 2000 x⸗x
        Courier New x⸗x
        Everson Mono x⸗x

        Junicode x⸗x
        RomanCyrillic x⸗x
        Segoe UI x⸗x
        Tahoma x⸗x
        Times New Roman x⸗x

    Thus, even if the users of the DOUBLE OBLIQUE HYPHEN (the character was intended for special transliterations of Coptic; therefore the font list above contains several fonts designed for linguistic use) would tolerate a horizontal glyph, they hardly will encounter one. They are served perfectly with the existing fonts, all supplying their preferred oblique glyph, whether they use the fonts designed for linguists, or the universal fonts for everyday use like Arial or Courier New.

2. More appropriate would be the question asked the other way round: "Do the users of the proposed DOUBLE HYPHEN (i.e. the double horizontal hyphen) have a problem if their character is displayed by an oblique glyph?"

    As outlined in the proposal, there are at least the following three main user groups for such a character:

    •   People who transcribe handwritten documents from the 19th century and earlier, or texts of that era printed in Fraktur, use the double hyphen (in its horizontal form) to preserve the hyphens in the original texts (where they are horizontal or oblique, the latter being common for printed texts), especially in

cases where the modern orthography requires compounding instead of hyphenating.

The purpose of such transcriptions is to present the appearance of the original texts in modern typographic manner, not to transform the texts according to modern orthographic rules, which would require a single hyphen for doing so. (Anyway, Unicode does not standardize orthographies, thus the fact that the modern German orthography as taught in schools does not use different hyphens is no counter-argument against encoding).

As outlined in the original proposal, the double oblique hyphen is considered not appropriate in modern typography, thus the horizontal double hyphen is the character that is used to transcribe hyphens of the original text.

- German literates and literature scientists, who want to edit, cite, or discuss the work of the German author Arno Schmidt (1914-1979), who used a non-standard orthography in a significant way.

- Individuals who adhere to this non-standard orthography or some of its constituents, especially using the double hyphen to emphasize the literally meanings of the constituents of German compound words.

  Here also, as Unicode does not standardize orthographies, the fact that the use of the double hyphen is not part of the orthography taught in German schools is no counter-argument against encoding.

All these three groups use the equals sign rather than the double oblique hyphen when they do not have access to a double (horizontal) hyphen proper (as it also is outlined in the original proposal), in spite of the fact that the equals sign is usually too long and therefore causes inferior typographic quality.

Thus, it is to be concluded that these users are not satisfied if their character is displayed by an oblique glyph when common modern typography is used.

3. Also, there is no obvious advantage in unifying the "double oblique hyphen" and the "double (horizontal) hyphen" into one character. (In fact, neither the PRI text nor the minutes of UTC #123 state such an advantage.)

   On the other hand, such a unification forces users to select a specific font to get the character they want. Especially, the users of the "double oblique hyphen", who now can depend that any font which supports their character at all will provide it correctly, will be confronted with a retrograde step when new fonts (or new versions of their fonts) start providing a glyph which they do not expect (especially when their rely on the Unicode rules of character identity).

4. Moreover, nowadays "universal fonts" are common which try to serve a part of the user community as large as possible, without giving preference to the special needs of any parts, and even without requiring the font designers to have even any knowledge of the special needs of special user groups. The designers rely solely on the Unicode standard itself.

   Especially, such "universal fonts" serve the needs of the Internet community, where the site designer usually cannot rely on the availability of special fonts on the user's side. Instead, the designer relies on the availability of such "universal fonts".

   Also, requesting the user to select an "appropriate" font to get their character to be displayed correctly is a fallback into the era when special fonts were needed to overcome the drawback of limiting the number of characters within a font by 8-bit codes, when you e.g. have to use a special IPA font which maps IPA characters on the position of "ordinary" characters in a proprietary way. Unicode was introduced to use to overcome this.

   To reintroduce such mappings like *»when you want a "double horizontal hyphen", use a special font where this is mapped onto the code position of the "double oblique hyphen"«,* is like revitalizing the old-fashioned *»when you want an IPA "ɛ", use a special font where this is mapped onto the code position of the "e"«.*

Thus, the only way to serve all user groups without unnecessary impediments, and in accordance with the Unicode principles, is to give the "double (horizontal) hyphen" a separate code point, as proposed.