

To: UTC
From: Debbie Anderson, Script Ad Hoc Chair
Date: 6 June 2023
Subject: Hungarian FORINT SIGN - Response from Vacek Nules

Below is further correspondence from the author of the Hungarian FORINT SIGN proposal ([L2/23-060R](#)). The author is responding to the report from discussion at UTC #175 that some members recommended users should use either F-ZWJ-t (<0046, 200D, 0074>) or Ft (<0046, 0074>).

From: Vacek Nules
Date: Saturday, April 29, 2023 at 12:12 PM
To: Debbie Anderson

I think a sequence is unacceptable because the computer systems supporting Ft signs still in productive use are DOS-based (accounting, billing, ticketing and POS systems etc.) which usually use naïve text consoles and are particularly picky about correspondence between one byte, one character cell and one character (older 8-bit ones are even more so, but they are unlikely to be in serious use), with no pretense for multi-byte sequences or zero width characters, hence a need for an atomic character.

For example, look at this file detailing Unicode mappings for FreeDOS codepages:
<https://github.com/dosbox-staging/dosbox-staging/blob/608ca8fb7869cc5e8c618c0e6204d15c2981cc92/contrib/resources/mapping/MAIN.TXT>

This lists codepage 3845 but not 57781 for Hungarian. The difference between the two is a single character: the forint sign, which was added in to FreeDOS around 2005 and had not yet propagated through into this project.

However, it lists Unicode mappings for most FreeDOS codepages, including several Polish ones with złoty signs, with their mappings looking like these:

0x9b 0xfffd #replacement for PLN SYMBOL, symbol not available in Unicode
 in other words, mapping the one-byte codepoint 0x9B to U+FFFD, for lack of a better atomic character, instead of a sequence of 007A, 0142 or 007A, 200D, 0142. This is because of the aforementioned character encoding model and would affect code page 57781 as well had it been listed in this mapping table.

Would it help if I got a few endorsements from the online FreeDOS community?

Follow-up email-1

From: Vacek Nules
Date: Saturday, April 29, 2023 at 2:18 PM
To: Debbie Anderson

Just a hypothetical example: porting and Unicode-ifying a DOS software printing bills/receipts to paper would make fixed-length rows wrap over if we chose to replace the atomic character with any sort of sequence, with a ZWJ probably becoming an unit-length tofu on the printout.

Follow-up email-2

From: Vacek Nules

Date: Saturday, April 29, 2023 at 3:51 PM

To: Debbie Anderson

I think adding an annotation like this could be useful in persuading those opposed:

U+1CCFA  FORINT SIGN

- for compatibility with legacy Hungarian code pages
- in running text U+0046 F U+0074 f is preferred

By the way, I find U+1CCFA an awkward codepoint, located in an empty space between two blocks of outlined alphanumerics which was clearly reserved only to make the codepoints of the outlined digits have the respective last digits. U+1CEBx would be better, but as it *is* a currency symbol I don't think the code point of U+20C1 is problematic in any way. There are already a few currency symbols there which are considered "legacy."

If squeezing this proposal through succeeds, a properly written and sourced (unlike 22-092) złoty sign one should follow, with that character having a similar level of permeation and an identical rationale to the forint sign. I'm willing to write that one as well if there's any interest. :)