

Subject:RE: Revised mapping table

From: Michel Suignard <michel@suignard.com>

Date:5/11/2012 8:31 AM

To:"csluqin@comp.polyu.edu.hk" <csluqin@comp.polyu.edu.hk>

CC:'chen-zhuang' <chenzh-zhuang@163.com>, xiaomwang2006 <xiaomwang2006@163.com>

Return-Path: <michel@suignard.com>

Dear Lu Qin

I have just integrated the data and fonts received from IRG and here are my findings:

- I received two fonts (one from Japan by email and one from Taiwan in person in Chiang Mai) after I received the IRG package from the IRG editors. These two fonts, although more recent, proved inadequate for production and I used instead the versions received from the package.

- Many of the provided fonts have names that suggest that they are Extension D fonts which is confusing, in the future they should use Extension E in their logical name and filename instead (Font from Vietnam is already OK).

- For the proposed character at index 7251 (sources are TE-3D36 and MAC-00078), the 'temporary' code point is being mentioned as 7DA2. However neither the Macau font (more on this later) nor the Taiwan font contain that character at that location. I found it in the Taiwan font at 7D9F and modified the mapping accordingly. Please advise which of the two is correct (font or spreadsheet).

- The font provided by Macau (MSCS file name: Macao_MSCS.irk.ttf) is not using the temporary code points mentioned in the spreadsheet and therefore had to be remapped by hand. Needless to say, I could do because Macau has only 48 entries and the provided font had about 60 glyphs making the investigation manageable. However the font is missing a proposed character: index 7888, source: MAC-00034, temporary code point: 801F. At minimum I need a font containing that missing character, and better a font that obeys the encoding schemes used by the other sources.

As you can imagine, the remapping for these large numbers of entries (over 6000) has to be scripted based on the spreadsheet data, therefore I have to assume that the spreadsheet and the fonts are in sync. The Macau font was totally incorrect so it was an easy find. However finding the mistake for the font from Taiwan was pure luck (because it also had a source reference from Macau for the same code point). This makes the whole data set suspicious and will therefore require a more thorough review of the whole set by IRG members.

The result is now available in the WG2 web site as n4383.pdf (as mentioned in the WG2 meeting #60 resolution M60.19) being part of the additional repertoire approved for inclusion in the next edition of 10646 (going for CD ballot). The character in error (for MAC-00034) is located at 2C940 and shows an erroneous glyph. I would appreciate an update for the Macau font at your earliest convenience as I can't produce the CD for 10646 4th edition until then. I will probably revise n4383.pdf as soon as I get an updated Macau font, but to simplify your review I won't touch the other fonts (I was able to work around for the Taiwan issue).

Best regards,

Michel