ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)

L2/04-400

The following text proposes to document visually the various sources information for all CJK Unified Ideographs. Currently only the CJK Unified Ideographs block (4E00-9FFF) and the CJK Unified Ideographs Extension A (3400-4DBF) are presented in the standard in a multiple column format showing the sources as visually available in the various East Asian local standards. The proposal is to add similar information for the CJK Unified Ideographs Extension B (20000-2A6DF) and any other CJK Unified Ideographs future extension.

This information could be made available either through a new Annex to ISO/IEC 10646 or a separate Technical Report. Past consideration about document size should not be a consideration given the fact that ISO/IEC 10646 is not delivered in paper format anymore.

Benefits

- That information would give access to visual information not easily available from East Asian local standards. Very often they are only available with explanatory text in their local writing languages and with various formats that make usage difficult. Furthermore these standards are often not easily available.
- This provides guidance for vendors that wish to implement fonts for each CJK countries or regions. More and more users and government entities are making strong requirements to follow precisely the shapes mandated by the local standard mapped into ISO/IEC 10646 as documented by the source references already documented by the standard. While this is easily verifiable for CJK Unified Ideographs collections presented in multiple columns it is not so for Extension B.

Resources

- It is expected that IRG could easily provide the chart information for Extension B, given that the mechanism is similar to what was done for the other CJK Unified blocks. The same could be done for Extension C when it is proposed for inclusion in the standard.
- The project editor is volunteering to author the textual component of a new Annex to ISO/IEC 10646 or a new Technical Report.

Finally, because the CJK Compatibility Ideographs are typically single sources or can only be used by several sources if all of them use the same glyph, there is no requirement to add such information for those ideographs. However, few CJK Unified Ideographs are located in the CJK Compatibility Ideograph collection: FA0E-FA0F, FA11, FA13-FA14, FA1F, FA21, FA23-FA24, and FA27-FA29. These characters should also be documented in multiple column format; this is already done for FA1F and FA23 in Annex P.