ISO/IEC JTC/1 SC/2 WG/2

Universal Multiple-Octet Coded Character Set (UCS)

Secretariat: ANSI

Title: ~~Beyond the request of JIS COMPATIBILITY IDEOGRAPHS~~ revised

Doc. Type: expert contribution

Source: Takayuki K. Sato -Japan

Project: 02:18.01

Status: To be considered at #39 WG2 meeting in Greece (SEP2000)

Date: 2000-06-01

Distribution: ISO/IEC JTC/1 SC/2 WG2

Reference: WG2 N2197, Resolution M38.11 WG2 N2142, N2196, N2221R, N2222R

Medium:

Japan proposes WG2 to open the discussion on the issues described in this document.

Issue-1: Definition of the COMPATIBILITY IDEOGRAPHS in contrast with the UNIFIED IDEOGRAPHS in somewhere ISO/IEC 10646 is necessary:

The more people believes the CJK UNIFICATION, the more difficulty to understand an existence of the COMPATIBILITY IDEOGRAPH, for non native user of the CJK IDEOGRAPHS in particular. Therefore it might be better to include an explanation of both IDEOGRAPHS within the ISO/IEC 10646.

One sample of the explanation is definitions in N2196.   But, it may not be necessary to be so. It may be other ways to explain. .

Definitions in N2196 are: (for reference)

Definition-1: CJK COMPATIBILITY IDEOGRAPH　　CJK ideograph that has to be unified with existing CJK UNIFIED IDEOGRAPH in principle. However, different code position is assigned for the CJK COMPATIBILITY IDEOGRAPH to guarantee "round trip conversion" between UCS and existing national standard(s) or industrial practices.
Definition-2: CJK UNIFIED IDEOGRAPH　　CJK ideograph that is independent ideograph from any other CJK UNIFIED IDEOGRAPH per the unification rule of the UCS.　The unification is done by the unification rule defined in the ISO/IEC 10646.

By the way, Japan strongly request WG2 to confirm the N2196 as a guidance rule of the COMPATIBILITY IDEOGRAPHS.

Issue-2: Existing CJK COMPATIBILITY IDEOGRAPHS.

Since both WG2 N2142 and N2196 are requesting "compatibility information" for the COMPATIBILITY IDEOGRAPHS.　However, existing CJK COMPATIBILITY IDEOGRAPHS (F900-FA2D) do not have such an information.

There might be better to amend ISO/IEC 10646-1 by the compatibility information for those CJK COMPATIBILITY IDEOGRAPHS.

Issue-3: Is annex P right place to amend the compatibility ideograph?

The compatibility information has two key components. One is the source standard and another is "corresponding CJK UNIFIED IDEOGRAPH (or it's extension).

The source standard information is just a couple of lines of text for a group of the COMPATIBILITY IDEOGRAPH, thus annex P might be a good place to fill in.　Besides, the correspondence data might make the annex P too large (and not practice to use).

It is better to consider locating the corresponding data onto separated place. Some of ideas are:

- To create new annex for this data　or
- Add parenthesized data to the character name

Issue-4:　Mapping information.

To avoid the mapping data inconsistency problem.　It is better that the mapping data should be defined within the source standard.　ISO/IEC 10646 does have only a data of the name of the source standard for the COMPATIBILITY IDEOGRAPHS. The mapping data is only necessary for the users of the source standard.　The other users are not expected to use the COMPATIBILITY IDEOGRAPHS in principle.

-----end-----