ISO/IEC JTC/1 SC/2 WG/2
Universal Multiple-Octet Coded Character Set (UCS)
Secretariat: ANSI

Title: Proposal: Additional guideline in the WG2 Principles and Procedures

Doc. Type: Experts Contribution

Source: Takayuki K. Sato   -Japan

Project:

Status: to be reviewed at #42 WG2 meeting

Date: 2002-05-13

Distribution: JTC1 SC2 WG2

References: n2352R, n2002

Number of pages: 2

This paper proposes to add one more guideline on to the "WG2 principles and procedures n2352R".

The addition is about the coding model change request of existing script(s), or adding new coding model onto the UCS.

In principle, the proposal requests that principles and procedures should clearly state that the coding model change of the existing script should not be allowed and adding new coding model at the time new script addition on to UCS should be reviewed very carefully, and addition of new model should be allowed only when the new model is the only way to code the script(s) to be added.

Out from the character addition/deletion requests for UCS, sometimes it has been observed that the request is actually a coding model change request.   Even though the request is going to share a code point(s) with same character shape printed, the behavior of the code point(s) are very different from the code point which is already available.   Submitters of the proposal usually

do not pay much attention on the behavior changes but it is actually very significant technical change of the code point. The change may cause in-compatibility data problem between old and new that should not happen for coded character set standards.

Avoiding this kind of problem by sticking on one model is a kind of common sense among the developers of the coded character sets. However, this is not a common sense for out side coded character set developers. Since WG2 is getting into much rear used characters coding, and need to deal with the out side people, it is better to state this point clear on to our principles and procedures.

In addition, minimizing a number of coding models is another very basic understanding among coded character set developers, but this is also not a common sense approach for some of the people. Thus, it is also better to state this on our principles and procedures.

Note that this is nothing new, just a clarification and reflection of WG2 common understanding onto the paper that to be written by outside WG2 people.

--end---

PS: "No duplicated encoding for same character" may be added also.