L2/02-133

ISO/IEC JTC/1 SC/2 WG/2 N2429

2002-04-01

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

Title: Proposal: Use full plane-13 for the Han variation selector
Doc. Type: Experts Contribution (also NB of Myanmar position)
Source: Takayuki K. Sato and Tatsuo L. Kobayashi -Japan
Project:
Status: to be reviewed at #42 WG2 meeting
Date: 2002-04-01
Distribution: JTC1 SC2 WG2
References:
Number of pages: 2

There is a variation selector proposed in plane-14. We reviewed a possibility of application and necessary number of the selector for a CJK unified ideographs.

We recognize that there are two kinds of possible application of variation selector for CJK unified ideographs.

a. To specify a variation of CJK unified ideograph that is already existing in UCS.

b. To specify a specific print and/or display shape of the ideographs. (like Math symbols)

Note that there are two kinds of usage of variation selector already expressed in WG2. One is to specify exact shape for Math symbols (this is the same as b above, and another is the case of Mongolian application. The Mongolian application is an "usage as a glyph selector". Thus the case a above is the third and new application of the variation selector.

1. Specifying a variation.

To specify a variation means that the user want to apply users own unification rule.

Note that the unified ideograph that is specified by the variation selector has its own acceptable micro variation range.

Therefore, even if the printed shape on the UCS with variation is the same, the allowable variation is different from person to person and or application to application.

As far as the necessary number of the variation for one kind of application. 256 might be enough, but the necessary number of the kind (individual, user group, application etc.) might be tremendous. At least 1024 is necessary per country.

2. Specifying a shape

We reviewed the necessary number of the shape variation in case of the application which need to specify exact shape.

One case study gave enough information. U-58FD (and U-5BFF) has more than 10,000 variation in shape and those are clearly distinct each other.

There is a book which shows a collection of 10,000 of the character (unfortunately only two in UCS)..

This means that the number of the variation selector for this application should be more than 10,000.

3. Proposal

To accommodate above nature of CJK unified ideographs. We propose to set up full VARIATION SELECTOR plane such as plane-3.

The plane should be many blocks. One block for more than 15,000 code points for specific shape application.

Others are more than 1,000 blocks with 512 code points each.

The necessary number of the blocks should be reviewed by CJK related countries, and if it is necessary, the necessary plane should be extended into two such as plane-3 and –4.

----end-----