

Subject:

Rewording I.2 B .. in WG2 p&p doc N3452

From:

Uma Umamaheswaran <umavs@ca.ibm.com>

Date:

Fri, 6 Nov 2009 16:26:06 -0500

To:

csluqin@comp.polyu.edu.hk

CC:

Mike Ksar <mikeksar@gmail.com>, Michel SUIGNARD <Michel@suignard.com>, kenw@sybase.com

Dr. Lu:

Below is a section from the WG2 principles and procedures document

www.dkuug.dk/jtc1/sc2/wg2/docs/n3452.pdf - Annex I on Guideline for handling of CJK ideograph unification and/or disunification error. Item B here is related to which glyph is to be selected for a retained ideograph when another one will be encoded under item A. The wording after the second sentence has to be revisited in light of the multiple-column charts that will be in the next edition as well as in next Unicode version that will be in synch.

I would appreciate it if you or some experts in IRG can suggest suitable rewording as to which glyph will remain etc. that will apply in the context of multiple-column charts everywhere. Not sure if IRG should have a discussion on this topic .. if so, please have it on your next meeting agenda and make some recommendations.

Thank you and best regards, Uma

I.2 Guideline for “to be disunified” errors

A. The ideographs to be disunified should be disunified and should be given separate code positions as soon as possible (disunification in some sense, and character name change in some sense also). These ideographs will have two separate glyphs and two separate code positions. One of these ideographs will

stay at its current encoded position. The other one will have a new glyph and a new code position.

B. For the ideographs that are encoded in the BMP, the code charts in ISO/IEC 10646 are presented in multiple columns, with possibly differing glyph shapes in each column. The question of which glyph shall be used for the currently encoded ideograph will be resolved as follows. In the interest of synchronization between ISO/IEC 10646 and the Unicode standard, the ideograph with the glyph shape that is similar to the glyph that is published in the “Unicode Charts” will continue to be associated with its current code position. For the ideographs outside the BMP, the glyph shape in ISO/IEC 10646 and the Unicode Charts are identical and will be used with its current code position.

C. The disunified ideograph will have a glyph that is different from the one that retains the current code position.

D. The net result will be an addition of new ideograph character and a correction and an additional entry to the source reference table.

V. S. UMAheshwaran, Ph.D.

Globalization Centre of Competency, IBM Toronto Lab

A2/SZ8, 8200 Warden Avenue, Markham, ON, Canada, L6G1C7; +1 905 413 3474;

Fax:905 413 4682; TieLine 313-3474; email: umavs@ca.ibm.com