INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 2/WG 2/IRG

Universal Multiple-Octet Coded Character Set
(UCS)

ISO/IEC JTC 1/SC 2/WG 2/IRG *N1644*
2009-11-26

Title:	**Suggested Revision of Text for Annex I of WG2 P&P**
Source:	**IRG P&P Drafting Group**
Action:	**For review by WG2**
Distribution:	**IRG Members and Ideographic Experts**
References:	

Annex I: Guideline for Handling of CJK Ideograph Unification or Dis-unification Errors

(Same as WG2 Principles and Procedures Annex I)
Source: www.dkuug.dk/jtc1/sc2/wg2/docs/principles.html

There are two kinds of errors that may be encountered related to coded CJK unified ideographs.

> Case 1: *to be unified* error - Ideographs that should have been unified are assigned separate code points.
>
> Case 2:*to be dis-unified* error - Ideographs that should not have been unified are unified and assigned a single code point.   An example of this is the request from TCA in document N2271.

When such errors are found, the following guidelines will be used by WG 2 to deal with them.

I.1 Guideline for "To Be Unified" Errors

A. The "*to be unified*" pair will be left dis-unified. Once a character is assigned a code position in the standard, it will not be removed from the standard.
B. If necessary, an additional note may be added to an appropriate section in the standard.

I.2 Guideline for "To Be Dis-unified" Errors

A. The ideographs to be dis-unified should be dis-unified and should be given separate code positions as soon as possible (dis-unification 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. The ideographs that are presented as multiple columns in the code charts of ISO/IEC 10646 may possibly have differing glyph shapes in each column.   The question of which glyph will be used for the currently encoded ideograph will be resolved in the interest of synchronization between ISO/IEC 10646 and the Unicode standard.   That is, if corresponding ideograph has single column in "Unicode Charts", 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 presented as single column in ISO/IEC 10646, the glyph shape in ISO/IEC 10646 and the Unicode Charts are identical and will be used with its current code position.
C. The dis-unified 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.

I.3 Discouragement of New Dis-unification Request

There is a possibility of "pure true dis-unification" request. This is almost like the new source code separation request.   This kind of request will not be accepted disregarding the reasoning behind. Key difference between "TO BE DIS-UNIFIED" and "WILL NOT BE DIS-UNIFIED is as follows.

> a. If character pair is non-cognate (meanings are different), that pair of characters is TO BE DIS-UNIFIED.
> b. If a character pair is cognate (means the same but different shape), that pair of characters WILL NOT BE DIS-UNIFIED.

Dis-unification request with reason of mis-application (over-application usually) of unification rule should NOT be accepted due to the principle in resolution M41.11.

Deleted: For t

Deleted: encoded in the BMP

Deleted: ,

Deleted: in

Deleted: are presented in multiple columns, with

Deleted: as follows.   I

Deleted: ,

Deleted: outside the BMP