Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRGN 1546

Date: 2009-4-10

Source:	IRG Ad Hoc
Title:	Additional section to Annex S

Below in accordance with IRG resolution M31.6 is the IRG proposed additional section to Annex S of ISO/IEC 10646: 2003. This is the second of two documents produced under the WG2 endorsement of IRG review and feedback on Annex S given in WG2 resolution M51.36.

S.5 Clarification of unification procedure

The examples provided in this annex **are not to be considered exhaustive**, and therefore it is necessary to be able to apply the rules and principles set forth in this annex to new situations. The rules, principles, and illustrations in this annex are descriptions of the calligraphic traditions and conventions used for CJK ideographs, not a set of mathematical transformations.

When applying the unification rules and principles, the following points should be considered:

- Non-cognate characters are not unified.
- Cognate characters, meaning that they share the same readings and meanings, are unified according to the table below:

	Actual Shape	
	Different	Exact Match
Same Abstract Shape	Unify	Unify
Different Abstract Shape	Do Not Unify	

S.5.1 Clarification of differences of actual shape

In accordance with the unification based on the analysis of components model of S.1.3 a pair of glyphs with different actual shape but the same abstract shape can be used as components to generate other pairs of glyphs in the same way. For example, given 兌•兑 have the same abstract shape then the following are also pairs with the same abstract shape:-

悅·悦, 捝·捝, 敓·敚, 梲·梲, 涗·涗, 稅·税, 脫·脱, 蛻·蜕, 說·説, 銳·鋭, and 閱·閲.

Although (皡•皡) have the same abstract shape, and their corresponding components (皋•皋) also have the same abstract shape (cp. 翱•翱), this does not imply that the corresponding top right hand components (白•自) also have the same abstract shape outside of this particular context. A relatively complex character is sometimes modified when written smaller as part of another character.

The following groups of ideographs shown below are further examples of differences of actual shape but with the same abstract shape when used as components:-

天•天,孝•孝,印•卬,賣•賣, 戌•戍,另•另•另,亏•亏,己•巳,义•叉,
丏•丐,莧•莧,幵•开,東•柬,臿•臿,羊•羊,兹•兹•茲,叵•叵,奄•奄,
臾•臾 ,盇•盍,畀•畀,那•那,尢•尢,匚•匸,朁•朁 ,犬•大,杀•杀,
豖•豕,叉•叉 ,丈•丈,丸•九,玉•王,曳•曳,岂•岂,巛•巛,戋•戋•戋,
底•底 ,貨•貨,卑•卑•卑,虽•虽,賴•賴,疐•疐,友•犮,臯•臯,采•采,
廉•廉,象•象,轂•轂,缶•缶,缶•缶,羕•羕,羨•羨,网•网,睿•睿,卧•臥,
匕•匕•七,凡•凡,朮•术,曰•日,市•巿,畫•畫,耒•耒,弱•弱,爫•爫,害•
害,勺•勺,次•次•次,蔑•蔑,与•与,唐•唐,冉•冉,寧•寧,匈•匈•匈,画•
画,具•具,鬲•鬲,灰•灰,華•華,叟•叟•叟,业•业,着•着,瓜•瓜,丩•丩,
艮•艮•艮,敖•敖,及•及,止•止,悳•悳,夆•夆 ,豪•豪,壳•壳,鼠•鼠•鼠,
疌•疌,尨•尨•尨,取•取,梟•梟,虍•虍,处•処,角•角,䍃•䍃•䍃 ,巢•巢,
丬•爿 ,微•微,产•产,门•门,會•會•會.

<u>S.5.2 Clarification of differences of abstract shape</u>

In accordance model of S.1.3 a pair of glyphs with different abstract shape that have more than one component can be used as components to generate at least one pair of glyphs with a difference of abstract shape. For example given 閒·間 have different abstract shape then it follows 月日 also have different abstract shape.

Adding like components to pairs with different abstract shapes usually, but not always, results in pairs of different abstract shape. Though as S.1.4.3 'Different structure of corresponding components' illustrates, usually adding like components to a pair of glyphs with different abstract shapes leads to pairs of different abstract shape, therefore for example because 閒·間 are of different abstract shape then the same is true for 僩·僴,嫺·嫻,憪·憫, etc. However because the higher nodes take precedence in the S.1.3 model then it is permitted for the new pair of glyph formed have the same abstract shape, for example though 口·厶 do not have the same abstract shape adding 月 to both gives 肙·肻 which do have the same abstract shape.

The following groups of ideographs are further examples of differences of abstract shape:-

肙•冐,麻•麻,嘗•尝,黨•党,兩•两,尋•寻,產•产,當•当,會•会,僉•佥,
喬•乔,壽•寿,巠•圣.

Date: 2009-06-16
Source: Japan
Title: Review comments on Annex S draft (IRG N1545, N1546)
Status: Input to the Annex S ad hoc group discussion
Requested Actions: For review
Pages: 2

WG2 requests IRG to finalize Annex S revision text by 2009-07-17 (WG2 M54.21) although its schedule is behind. Japan expects this editorial group needs to be agreed within this meeting period and any text which does not reach to agreement should be dropped for the future revision.

This document describes Japan's comments on IRG N1545 and IRG N1546.

Comments on IRG N1545 (Changes to Annex S)

Japan basically thinks IRG N1545 is OK to submit to WG2, however some minor modifications are needed.

The word "L-shape" in (1) does not point Ideographs, but rather IDC character. However, the word is not defined as a term, therefore it may be misunderstanding. It should be explicitly denoted or be replaced with other words.

(Editorial) Title of S.1.4.3 should be "Different Structure of a corresponding components." "Different number of components" is a title for S.1.4.1.

"Non-unifiable" issues are common to S.1.4.1 to S.1.4.3, but it only emphasize on S.1.4.3. This statement is not necessary.

(Editorial) At the first sentence of S.4, "in S.1. The pairs" should be "in S.1. of this annex the pairs".

Comments on IRG N1546 (Additional section to Annex S)

Japan considers that the statement of Annex S.5 is structurally difficult to understand, somewhat repetitive of previous sections, and does not provide meaningful information to readers. As of it, Japan thinks that the entire S.5. should not be attached to Annex S at all.

For example, S.5.1 can be considered as mere re-statement of S.1.5. Examples appeared in S.5.2 is also a mere repetition of S.1.4. The statement of S.5.2 tells some natures on abstract shapes, but does not provide any meaningful rules on them, so it is not useful to common readers.

Paragraph below the title do not constitute any meaningful statement.

Japan thinks it is unnecessary to use chart in S.5. S.1.2 can be read as “in case of variants can not be unified, their abstract shapes are different” and this is enough.

The example of 「兌」 case does not lead to any conclusion. It is indeed true that same abstract shape may contain different component, but that does not lead to any meaningful conclusion. Examples are hardly useful to any kind of readers.

S.5.1, page 2, first paragraph

The following groups of ideographs shown below are further examples of different actual shape, but when used as components of cognate characters, are considered as the same abstract shape.

For the examples shown in S.5.2, the statement does not seem to reflect what the statement trying to say.

It is impossible to exhaustively enumerate all “non-unifiable” examples, so any glyphs should not be listed here.

This entire document should be removed from Annex S, as they are misleading and unnecessary.

(End of document)