

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N1420

Date: 2008-06-02

Source:	Ad Hoc Chief Editor
Title:	Annex S Review Ad Hoc (2)
Actions required	Review and discussion
Distribution:	IRG Members and Ideographic Experts
Medium :	Electronic

The main focus of the Ad between IRG #29 and IRG #30 was on the clarification of further examples of CJKV components that have either the same or different abstract shape. Comments both brief and detailed were received from various Ad hoc members.

Examples have been included because there are of interest to one or more members of the ad hoc. Among other things they have been chosen to clarify simplified vs non-simplified components, and the data in <http://kanji-database.sourceforge.net/housetsu.html>.

Some discussion of the overall content of Annex S took place regarding which can be summarised briefly as :-

- 1) that changes to the existing Annex S should be kept to a minimum.
- 2) that there is a need for further examples and information to be added to Annex S
- 3) that some of the additional examples are different in type to some of the existing basic examples.

In view of this the suggestion of the Editor is results of the review process and IRG discussion be split into at least two documents (1) requested changes to the existing content of Annex S (2) requested additions to Annex S.

Thanks in particular to Retarkgo Yan and Taichi Kawabata for their many suggestions and corrections, too numerous to mention item by item.

Matters requiring further discussion

(including a summary of comments received).

Appendix A part one.

These are proposed groups of existing characters that illustrate what it is to have the same abstract shape. Though a few of these examples were looked at briefly during IRG #29, the whole list needs to be discussed in detail. The non-cognate status of a few examples was suggested, and the connection of example 52 with two characters under review noted. Though no objections to any examples was voiced in the ad hoc itself, some discussions outside of the ad hoc suggest whilst most are widely accepted that one or more of the examples do not have automatic support.

It is suggested the final list start as follows:-

"The following groups of ideographs shown below are further examples of differences of actual shape but with the same abstract shape, that are for various reasons separately encoded.

天	夭	孝	季	印	邛
5929	592D	5B5D	21949	536C	536D

..."

Appendix A part two.

The groups of glyphs here were agreed in principle at IRG #29. Only two examples received comments. It was noted that the glyphs in example 99 are encoded separately and therefore it is suggested this example should be moved to section A part one. Also there was discussion about the comment added to example 78, the nature of this example needs to be clarified. It is worth noting that equivalence of abstract shape is independent of meaning. Therefore example 78 should be discussed again.

It is suggested the final list start as follows:-

"The following groups of ideographs shown below are further examples of differences of actual shape but with the same abstract shape.

耒·耒, 弱·弱, 冫·冫, 害·害, 勺·勺, 次·次·次, 蔑·蔑,

..."

Appendix B part two

Though agreed in principle at IRG #29 these attracted a number of comments both general and specific.

General comments centred around:-

(1) Are all the examples necessary?

(2) the examples are more complex than those in section 1.4 of Annex S.

It is noted (1) that all the examples chosen are considered important by at least one ad hoc member and that this is sufficient for inclusion (2) the examples are in many cases more complex than those found in section 1.4, and so these new examples should not be placed in section 1.4 .

It is suggested the "reason" column either be removed altogether from Appendix B, or in any examples requiring clarification be replaced by one or more of 1.4.1,1.4.2 or 1.4.3.

Specific comments included:-

(1) that exceptions to examples 133, 134 and 135 do exist in the standard.

(2) alternatives representative glyphs for examples 131, 134 and 135 suggested.

Therefore 131,133, 134 and 135 should be reviewed in detail at IRG #30

It is suggested the final list start with:-

"The following groups of ideographs shown below are further examples of differences of abstract shape."

followed by either

"土·工, 胃·冒, 稟·稟, 然·焚, 戩·戩, 麻·麻, 愛·爱, 罷·罢,

..."

or

"

土	工	胃	冒	稟	稟
571F	5DE5	5190	5192	7980	7A1F

..."

Appendix E

This includes items postponed from IRG #29 for further discussion and also new examples raised during the ad hoc. Many of the items have not been reviewed so far by the IRG. All the items included, whether commented on or not, need to be discussed in detail, comments received where possible have been included in the document.

It is suggested that the examples agreed upon be added to appropriate lists above.

IRGN1400

This is the revised version of IRGN1387. whilst discussed at length by a small number of people, comments are invited from all editors.

IRGN1397

A number of IRG #29 raised Annex S issues and the results of discussions of these summarised in IRGN1397. This include a number of changes and additions not mentioned in the documents above, this should be reviewed and those agreed upon be included in the request for changes and additions to Annex S.

Conclusion

Whilst this ad hoc will hopefully result in many improvements to Annex S, it is also noted that the need to extend Annex S is as open ended as the the work of the IRG itself . Considered should be given to different options for achieving this aim efficiently.