

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1178

Date: 2005-11-30

Doc. Type:	Member body contribution
Title:	Recording of CJK(+A&B) Errors
Source:	CJK_C1 editorial group
Status:	Input to IRG
Action:	To be reported to the plenary
Distribution:	IRG Members and Ideographic Experts
Attached files:	1. Recording Format of CJK(+A&B) Errors 2. IRG N1172
No. of pages:	
Medium:	Electronic

The CJK_C1 editorial group met on Nov. 30, 2005, at IRG #25. Editors are from China, Japan, R.O.Korea, Vietnam, USA, Hong Kong SAR, TCA and Unicode Consortium.

The editors agreed to keep and update an internal list of errors found in CJK, CJK_A and CJK_B is necessary. Editors also agreed:

- The IRG internal CJK(+A&B) error list should contain not only the errors confirmed by editors but also the original error reports from the editors, as well as the detailed reasons for the results of discussion among the editors.
- The IRG internal CJK(+A&B) error list should be updated dynamically according to the newly submitted error reports and discussion results made by the editors.
- The IRG chief editor is responsible for maintaining the IRG internal CJK(+A&B) error list.
- A concise error list based on the IRG internal CJK(+A&B) error list should be generated and submitted to WG2 by the IRG chief editor when it is required by WG2 resolution or IRG resolution.

The editors reviewed and discussed a draft recording format for CJK(+A&B) errors (IRG N1164), and agreed to submit error reports and maintain the IRG internal CJK(+A&B) error list with the modified format. For detail of the modified format, read the attachment 1 please.

Attachment: Recording Format of CJK(+A&B) Errors

The recording format described in this document should be used for both of the IRG editors' CJK(+A&B) reports and The IRG internal CJK(+A&B) error list. It can be improved later if IRG editors require so.

For table use:

- When editors are submitting reports relating to characters duplicated or might be unified, Table 1 should be used.
- When editors are submitting reports relating to characters might be dis-unified, Table 2 should be used.
- When editors are submitting reports relating to glyphs with errors, Table 3 should be used.

Table 1: for characters duplicated or should be unified

Example:

Serial No.	UCS code point 1	Glyph 1 in ISO/IEC 10646: 2003	Source References	UCS code point 2	Glyph 2 in ISO/IEC 10646: 2003	Source References	Rationale provided by submitter	Conclusion (Y or N) drawn by all editors
xxxx-U-0001	219B4	𪛗	G_HZ	219AD	𪛗	T7-4D52	Because, see Annex S.... .., etc.	Yes. Reason:
xxxx-U-0002	27729	𪛘	G_KX T7-3B5F	04666	𪛘 𪛘		Because, see Annex S.... .., etc.	Yes. Reason:
xxxx-U-0003								

Since some glyphs had been corrected in *ISO/IEC 10646: 2003*, i.e. may be different to the original glyphs in *SuperCJK 14.0*, glyphs in *ISO/IEC 10646: 2003* should be used in the table.

The UCS code point 1 should be larger than the UCS code point2. It means that the character in the right was encoded

before the left one, therefore, the left one should be unified to it.

The source references in this table are optional. They are used to make the IRG members of the sources pay more attention to their characters. They can be found in the CJKC_SR.txt and CJKU_SR.txt in the CDROM of *ISO/IEC 10646: 2003*.

Table 2: for characters should be dis-unified

Example:


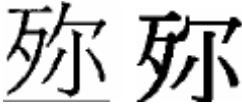
Serial no.	UCS code point	Glyph in ISO/IEC 10646: 2003	Glyphs should be dis-unified.	Glyphs are from	Rationale provided by submitter	Conclusion (Y or N) drawn by all editors
xxxx-D-0001	239E2	 239E2		Super CJK 14, 239E2	Because, see Annex S.... .., etc.	Yes. Reason:
xxxx-D-0002						

Table 3: for glyphs with errors

Example:

Serial no.	UCS code point	Current glyph in ISO/IEC 10646: 2003	Suggested glyph	Preceding Glyphs are from	Rationale provided by submitter	Conclusion (Y or N) drawn by all editors
xxxx-G-0001	25016	𪛗	𪛗	Super CJK 14, 25016		Yes. Reason:
xxxx-G-0002						

Note:

- The error report submitter should request for a document number at least 1 week before submission and make sure the document number is given by the IRG Rapporteur.
- “xxxx” in column “serial No.” means IRG Nxxxx of the original error report from the submitter.
- Every serial number in the above 3 tables should be unique, it should be kept with no change in the updating IRG internal CJK(+A&B) error list.
- The column “Conclusion (Y or N) drawn by all editors” should be kept empty when editors are submitting their error reports. Other columns in the table should be filled by editors.

- The column “Rationale provided by submitter” should be filled in detail as much as possible, such as “according to Annex S, compatibility ideographs”, etc., not “they should be unified... ..” or something else only.
- The submitted error reports should be Microsoft Word documents.

Caution: The UCS code points in tables may not match glyph images by mistake. Edit and review error reports carefully.