

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2/IRG N1357

Date: 2007-11-12

Source:	John Knightley
Title:	Presenting Evidence
Status :	For discussion
Distribution:	IRG members and Ideographic Experts
Medium :	Electronic

Presenting information and evidence about proposed new characters in a full and clear way, is essential for the IRG to carry out it's work efficiently. Whilst there exists well established standards and conventions concerning information about the shape and structure of characters (radical, stroke count, first stroke, IDS, etc.), for many other types of information and evidences (source, reading, meaning, variants, unification candidates, etc) both the presentation, and even the inclusion or otherwise of certain items has varied considerably for submission to submission. This document does not aim to answer all of the questions about this matter, but to be a step along the way of this process.

IRGN1347 6 characters and IRGN1367 120 characters both provided very full information about their respective character sets, and mark a breakthrough by being the first proposals to actively tackle potential unification issues. This noteworthy new feature is examined below.

Inclusion of Unification Candidates

The inclusion of unification candidates was both innovative and useful. In IRGN1367 this was included in the main table in the similar column.

(IRGN1367_J-urgent.pdf top of page 4)

#	j_source	Proposed glyph	IDS	Similar Shape	Variant Of	Evidences				FS	f	KX Rad	Strk
						Ref.	#pg	Source	Example				
32	JH-JTAF09	奏	𠂇尖天 Shape		奏 U+594F	Ref-2	14	新大字典		4	252.131	大	7
33	JH-IB0654	奥	𠂇𠂇𠂇𠂇米大	奥 U+5965	奥 U+5965	Ref-2	15	新大字典		3	253.211	大	10

[illegible]

Though not every proposed character has good unification candidates, these are the characters that require the most work to process. Furthermore considering potential unification candidates should be a core part of the process of preparing a set of characters for submission.

'Or other helpful information'

In IRGN1347-Appendix.pdf much 'other help information' including but not restricted to:-

(1) Japanese On and Kun readings for both the proposed character and the good unification candidate.

(2) Various readings and definitions.

(3) Multiple unification candidates.

(4) Direct reference to Annex S.

(5) Reference to UCS source glyphs.

(6) Information on relevant characters in proposed ideographic variation sequences

Conclusion

The inclusion of full evidence and other helpful information is essential for the efficient processing of proposed characters. Putting appropriate important information into a table form for viewing, and making all the text parts of that available in a suitable machine readable format, will increase the speed and efficiency of the review process by IRG editors. As in the areas of information about the shape and structure of characters, after due consideration standards need to be set and good practice encouraged.