

Universal Multiple-Octet Coded Character Set
UCS

ISO/IEC JTC1/SC2/WG2/IRG N2478Summary

Date: 2021-09-08

Source:	Chen Zhuang (The IRG Chief Editor)
Title:	IRGN2447 Response From China
Meeting:	IRG #57
Status:	
Actions required:	To be considered by IRG
Distribution:	IRG
Medium:	Electronic
Page:	2
Appendix:	

The author quickly reviewed IRGN2478 WS2021v1.0 including the Consolidated Comments, and consulted the Online Review Tools (ORT) Manager (Mr. Henry Chan), then presents his personal opinion in this document.

1. Total 4,925 characters

China: 1,223 (GDM 151, GKJ 768, GXM 4, GZ 300)

ROK: 191

SAT: 383

TCA: 1,000

UK: 1,000

UTC 153

Vietnam: 1,002

Note: The IDS checking shows 99 pairs and 1 triplet of identical IDS, which can be trivially unified, so actual count will be around 4,800 characters.

2. Acceptable Quality

1) Less than 7.31% Unifications

The comments raised around 309 unification pairs/triplets (618+ characters) including 129 WS2021 internal sharings (258+ characters). Therefore, less than 360 (618-258=360) of 4295, or 7.31%, are raised to be compared with coded characters for consideration of unification. This will be lowered when IRG discussion carries out.

2) Better evidences than before

The comments raised around 443 evidences while most are additional or better evidences thus acceptable, some of the others are debatable.

A few comments are debating on recognition of original glyphs or interpretation of original texts which may resulting changing character shapes.

A few evidences are questioned for their validity rather than quality. For instance, SN 04547 UK-20791 𠄎鬼肆, a character written as charm (符籙) in Taoist books. The comment is “IRG PnP 2.1.1.c says: characters must be used in script as characters in text. Logos and images used separately from running text are not acceptable.”

Note: Radical debatings are about understandings of characters, IDses, Glyph designs or glyph normalizations, Stroke counts, First strokes, T/S flags and others are considered minor issues, by view of the author of this summary.

3. Suggestion

Having considered the aboves, the author recommends IRG to process all 4,925 characters as a whole working set, although it exceeds the limitation of 4,000 ideographs (IRG PNP 2.2.1.b.1).