

Universal Multiple-Octet Coded Character Set International Organization for Standardization

Doc Type: Working Group Document

Title: IRG 41 Liason Report

Author: Dr. Ken Lunde, Unicode Liason & L2 Delegation Head

Date: 2013-12-11

IRG 41 took place in Tokyo, Japan from November 18th through 21st, and was hosted by IPSJ (*Information Processing Society of Japan*), ITSCJ (*Information Technology Standards Commission of Japan*), and JISC (*Japanese Industrial Standards Committee*). The three main topics were to make progress on Extension F1 (the first half of Extension F), finalize the new version of the PnP, and to go through the UNC (*Urgently Needed Character*) submissions. The IRG 41 Resolutions are now available.*

In addition to Michel Suignard, Lee Collins, and myself, who were L2/Unicode delegates, there were delegates from China (11), Hong Kong SAR (1), Taiwan (4), South Korea (2), Japan (8), and Vietnam (1), along with three individuals who happened to come from Japan (SAT).

The following is the future IRG meeting schedule:

IRG 42: Qingdao, China, 2014-05-19 through 2014-05-23 (approved by WG2)

IRG 43: San Jose, CA, USA (Adobe), 2014-11-17 through 2014-11-21 (approved by WG2)

IRG 44: Seeking host

New kIRG_ZSource

A new kIRG_ZSource was proposed to accommodate submissions that cannot be tied to a particular country or region, or which span more than one country or region, and which doesn't use UTC's kIRG_USource. Michel expressed reservations about this.

Urgently Needed Character Submissions

Three UNC proposals were discussed: IRG N1936 (UTC, 18 characters), IRG N1954 (UC Berkeley's SEI, 20 characters), and IRG N1967 (China, three characters).

IRG N1936, which was first presented during IRG 40, was discussed first. John Jenkins prepared a revised submission (2013-11-08) that incorporated feedback and suggestions from Japan, which also involved removing a controversial character (UTC-00955), which reduced the submission from 19 to 18 characters. Of the 18 characters in this submission, the most controversial one was clearly UTC-00791,† which is shown below, rendered in Kozuka Gothic (小塚ゴシック) in all six of its weights:



Japan and South Korea felt that the characters, as a whole, are not urgently needed, and the Rapporteur somewhat strongly objects to encoding UTC-00791. South Korea also made a statement to the effect that UNC should not be treated as an *express lane*. The characters that correspond to chemical element names gained the most favor.

* <http://appsrv.cse.cuhk.edu.hk/~irg/irg41/IRGN1980Resolutions.doc>

† <http://en.wikipedia.org/wiki/Biang>

In response, John Jenkins—being channeled through me—pointed out that some of these characters were delayed from Extension C, that the submitter should be given the benefit of the doubt with regard to urgency, and that the whole purpose of the UNC process is to encode smaller sets more quickly.

IRG N1974 included the following summary for this submission:

The Editorial group has reviewed the revised proposal from UTC IRGN 1936R(2013-11-13) with feedback from Japan and ROK. Members agree that some of the characters may qualify as urgently needed characters. Members do recognize that the character are currently being used, but some members still question the urgency for the encoding of these characters. As characters are produced and used over the internet is quite new, members still have questions about their stability and would like more time to consider it. Members are requested to provide their review and feedback for discussion in IRG Meeting No. 42. The UTC is also recommended to revise/consolidate the information for better readability.

John will prepare a revised version of the submission, which will be discussed during IRG 42. Andrew West also suggested stripping down the proposal to the bare minimum, which should increase its chances of being accepted during IRG42.

I introduced IRG N1954, which is a submission from UC Berkeley's SEI (*Script Encoding Initiative*) that includes 20 characters. Just prior to the introduction, the main author of the submission, Yuri Shardt, prepared, at my request, a statement of urgent need, which I found was missing from the submission. The most substantive comments came from Ken Chen (China) who observed that the left-right arrangement of the components for almost all of the characters are reversed compared to their reading order. I asked Yuri to prepare a statement that provides background material to explain this apparent discrepancy, which was posted as Appendix B, and which satisfied Mr. Chen. Japan felt that these characters are not urgently needed.

IRG N1974 included the following summary for this submission, which I relayed to Yuri:

The Editorial group has briefly reviewed the submission on Slavonic Transcription IRGN 1954. Members requested more time to review it as the nature of these characters, as a set, seems to be quite different from that of the other CJK characters. Some members also have reservation on the request to sort the characters using phonetic order rather than the CJK radical orders. The reason is that there are some transcription characters already coded in the CJK blocks. So, the phonetic order cannot be supported using UCS code. Sorting them different from radical order would only introduce more confusion.

This submission will be discussed in greater detail during IRG 42.

IRG N1967, which is China's modest UNC submission that includes only three characters, was discussed, and the consensus was that it was urgent. The IRG resolved to accept the submission, which means that it now goes to WG2 for approval. Interestingly, Andrew West found an article (in Chinese) about the first character in this UNC submission, which helps to bolster its urgency.‡

Extension F

Extension F was split almost down the middle by indexing radical into F1 and F2, and much of the time was spent on Extension F1 editorial work. This included a lot of questionable evidence from Korea, along with Japan wanting unification/disunification decision (so that they can handle unified characters via the IVD).

I carefully reviewed China's Extension F glyph change proposal (IRG N1968), and found that six of supposedly corrected glyphs do not adhere to China glyph conventions, which suggests that they pieced the glyphs together using parts from a Traditional Chinese font. Sequence numbers 26, 61, and 88 use the TCA-style 女, sequence number 82 uses the TCA-style 月, sequence number 83 uses TCA-style 公, and sequence number 87 uses TCA-style 今. Mr. Chen confirmed that these were genuine errors, and will be corrected.

Strokes

Related to Extension F1 editorial work, there was a lengthy discussion about counting strokes, specifically about whether single-source characters should use the actual stroke count, or the by-the-book stroke count. Given that horizontal extensions are possible, which I pointed out, the consensus was that by-the-book stroke counts should be used. Michel indicated that the *kRSUnicode* field allows additional radical/stroke information.

‡ <http://news.thmz.com/col89/2010/01/2010-01-08682302.html>

IVD

Although not discussed by a particular document, I observed that “IVS” and “IVD” are no longer being treated as four-letter words in the IRG, in part because Japan has begun to embrace the IVD as a viable solution for their large number of (mostly unifiable) variants.

National Standards & Horizontal Extensions

My document about continued national standard development and horizontal extensions (IRG N1964) was discussed at length. I think that I got my point across, but like the UTC, the IRG is not convinced that the *kIRG_{G,T,H,M,J,K,KP,V,U}Source* references are the best mechanism for conveying this information. Hong Kong’s activity report indicated that 553 additional characters were accepted for Hong Kong use, and my document prompted them to better document these additions. Japan spoke of using ISO/IEC 10646 Annex A for this purpose, but some people expressed their doubts about its practical usefulness. No action was taken at this time.

Other Tidbits

Bear Tseng quit TCA (Taiwan).

Hong Kong SAR identified 553 additional characters for Hong Kong use, which are in ISO/IEC 10646, but outside the scope of Hong Kong SCS-2008.

Toshiya SUZUKI (Japan) continues to do an absolutely superb job at representing Japan for all matters pertaining to the IRG. Because his organization does not support his IRG-related work, he tends to attend only IRG meetings that are held in Japan, which is incredibly unfortunate.

Version 6.0 of the IRG’s PnP (*Principles & Procedures*) document was finalized as IRG N1975.

That is all.