

Universal Multiple-Octet Coded Character Set International Organization for Standardization

Doc Type: Working Group Document

Title: IRG #47 Liaison Report

Source: Dr. Ken Lunde, Adobe Systems Incorporated

Date: 2016-10-23

IRG #47 took place in Matsuyama, Shikoku, Japan from 2016-10-18 through 2016-10-21, was hosted by METI and IPSJ/ITSCJ (on behalf of JISC), and was held at the Matsuyama Multi-Purpose Community Center. The main discussion topic was IRG Working Set 2015 Version 2.0 (Extension G), in terms of handling comments and feedback from the latest review cycle. The [IRG #47 Recommendations](#) (aka IRG N2180) are now available.

Including myself as the sole US/Unicode representative, in attendance were 32 representatives and experts from China (12), Hong Kong SAR (1), TCA (1), ROK (5), Japan (10), SAT (1), and one invited expert from Hong Kong SAR.

As stated in Recommendation IRG M47.1, the following is the future IRG meeting schedule:

IRG #48: Bundang, ROK, 2017-06-19 through 2017-06-23

IRG #49: San José, California, USA, 2017-10-16 through 2017-10-20

IRG #50: Macao SAR, 2018-05-21 through 2018-05-25 (tentative; China is backup)

IRG #51: China, 2018-10-22 through 2016-10-26 (tentative; Hong Kong SAR is backup)

Extension F Status

Extension F was briefly discussed, and the good news is that no changes were made. I did mention the UTC's request for *kTotalStrokes* data for Extension F, which is needed for Unicode Version 10.0, and it seems that I need to work with China offline to get that data. Worst case, I will use the *kRSUnicode* data to synthesize it. (Note that the requirement to include *kTotalStrokes* data as part of submissions was added to the IRG's P&P after Extension F.)

Extension G Status

The vast majority of the meeting was spent performing Extension G editorial work. Hong Kong SAR performed a quality check of the review comments, which indicated that the UTC's review was considered to be very good.

Anyway, the following UTC-submitted characters were either unified or withdrawn:

00003 (UTC-00969) Unified with 00005 (U; SAT-submitted)

01182 (UTC-01005) Unified with U+5F50 (U)

02151 (UTC-01258) Withdrawn (N; do not unify with U+2C27A)

04411 (UTC-01297) Unified with 04416 (UTC-01299)

04941 (UTC-01310) Withdrawn (W)

05125 (UTC-01311) Withdrawn (W)

A large number of UTC-submitted characters needed better evidence, which was supplied for all but two, which are listed above as withdrawn. The following two characters were at risk of being either unified or rejected, but were successfully defended:

03560 (UTC-01286)

04319 (UTC-00791)

In terms of UK-submitted characters with U-Source references, I had several exchanges with Andrew West during IRG #47, which involved reporting the status of particular characters that were at risk of being unified. His timely replies were very helpful.

Per Recommendation IRG M47.2, the following is the Extension G working schedule between now and IRG #48:

2016-11-18 IRG experts submit updated bitmaps to the IRG Chief Editor, and provide supplementary information for any conditionally accepted characters to other IRG experts

2016-11-25 IRG experts distribute questions (if any) to the supplementary information

2016-12-02 The IRG Chief Editor distributes IRG N2179 (aka IRG Working Set 2015 v3.0)

2017-02-24 IRG experts submit review comments to the IRG Chief Editor

2017-03-10 The IRG Chief Editor distributes consolidated comments

2017-04-28 IRG experts submit to the IRG Rapporteur responses to the consolidated comments

IRG's P&P & IVD

Version 9.0 of the IRG's P&P (*Principles & Procedures*) document was approved, and includes a change to Section 2.1.3 (*Non-cognate Rule*), along with the following new statement in Section 2.2.3 (*Required Evidence to be Submitted*): ***The appearance of a character as a head entry in a dictionary is generally considered evidence of actual use if the dictionary is listed in Annex D or is otherwise accepted by IRG as an authoritative source.*** (I had a couple of productive email exchanges with Andrew West that led to a version that was acceptable to the IRG.)

Speaking of the IRG's P&P, per Recommendation IRG M47.7, the following statement is very encouraging: ***IRG would also like to consider the IVD for handling CJK variants in the future, especially for IRG Working Set 2017. Experts are encouraged to submit their proposals for IVD registration in future IRG work.*** There is a 2017-05-19 deadline associated with this.

With zero prompting nor encouragement from me, the IRG openly discussed the handling of unifiable variants via the IVD, which is why it was so prominently mentioned in Recommendation M47.7. Years ago, we asked the IRG to include a statement about IVD use in their P&P, so it is excellent news that this is finally going to happen. I plan to draft a proposal to be discussed at IRG #48, which will first be discussed at UTC #150 or #151.

Glyph Normalization & UCV/NUC

Due to discrepancies between the representative glyphs that ROK has submitted and their

evidence, the IRG requested a meeting or two ago that they submit a document that formally describes their glyph normalization rules (see [IRG N2154](#)). Per Recommendation IRG M47.5, ROK will revise the document to make a distinction between normalization and unification, the latter of which are to be handled via the IVD.

Henry Chan, a student from Hong Kong SAR who was an invited expert and who represents the next generation of IRG experts, submitted several T- and G-Source error reports, along with suggested additions to the UCV (*Unifiable Component Variations*) and NUC (*Non-Unifiable Component*) lists, which are used as examples to guide unification and disunification by illustrating pre-existing cases. Per Recommendation IRG M47.3, many of Henry's UCV/NUC suggestions were accepted, and some still need to be discussed. I also noted that some of the error reports may result in orphaned Extension B characters. Per Recommendation IRG M47.6, China and TCA need to respond to these error reports.

IRG Working Set 2017

Per Recommendation IRG M47.8, the IRG requests that national bodies provide preliminary information about the characters that they plan to submit for IRG Working Set 2017 (aka Extension H), which will be discussed during IRG #48 in order to estimate the workload..

That is all.