LU, Qin [COMP]

From:	Michel Suignard <michel@suignard.com>
Sent:	Sunday, 15 July 2018 6:21 AM
To:	chen zhuang; Satoshi YAMAMOTO
Cc:	Lu Qin
Subject:	RE: Re:Questionable updates on revised CJK G code table regarding IRG#50 editorial report

Concerning my own action down:

7.2
#00731 u+3027E Ambiguous record (About PDAM2.2 only) UK source is not added.
> Michel, please add UTC-01211. (This not UK, it is in the excel file.)
#00854 u+302F6 Ambiguous record (About PDAM2.2 only) UK source is not added.
> Michel, please add UK-01357. (This is in the excel file.)

As of now, all 'U' sources are merged in one single field. In other words, we have not allowed the U sources to have multiple values (UTC, UK and USAT). I could possibly create records as follows:
U+3027E   kIRG_USource   USAT-06429 UTC-01211
U+3027E   kRSUnicode   32.8

The caveat is that you only have one representative glyph for all 'U' sources. If that is acceptable, I could implement this for these 2 characters and extend this for cases in the future where we have self-restricted ourselves.
The following screen shot shows the result for U+3027E (apparently the code can only deals with 2 sources, but that could be updated in the future).
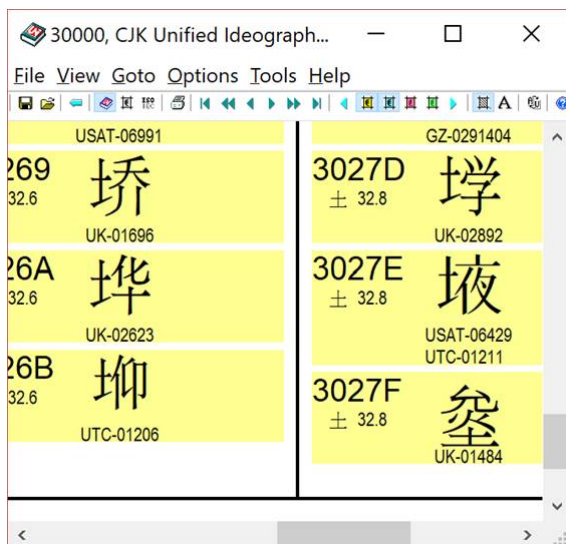
Having multiple glyphs for U sources would mean creating more kIRG_xSource fields, which is a different discussion. This would require significant change in the tool (Unibook), not a short-term option.

Concerning collecting comments and process in general: I can implement modest changes before going to ballot next month (but nothing like changing RS values which would reshuffle the repertoire). Glyphs change would be OK (assuming I get a font, although I would probably just use the updated glyphs and insert into existing to minimize regression risk). I can make the ballot period short enough that the CD ballot closes just before Hanoi, and then we can use Hanoi to analyze comments and come up with consensus among IRG members.

Best regards

Michel

From: chen zhuang <chenzh-zhuang@163.com>
Sent: Saturday, July 14, 2018 2:32 AM
To: Satoshi YAMAMOTO <yama.motch@gmail.com>
Cc: Lu Qin <csluqin@comp.polyu.edu.hk>; Michel Suignard <michel@suignard.com>
Subject: Re:Questionable updates on revised CJK G code table regarding IRG#50 editorial report

Dear Yamamoto san and Michel and Dr. Lu,

I fixed some data and marked them in blue as below, the red entries are for China, Korea and Michel to respond. I can fix all of them in the data file (excel) when I received feedback from China and Korea. I will trasfer the red entries to China and Korea.

For the matter of review, We may need one mor round review, but can we follow up the ballot schedule? Do we need to confirm the ballot comments in IRG51 (Hanoi, Vietnam 2018-10-22/26)?

Best regards,

Chen Zhuang

在2018年07月14 04时12分, "Satoshi Yamamoto"<yama.motch@gmail.com>写道:

> Dear Dr. Lu and Chen Zhuang,   cc: WG2 Project editor
>
> As mentioned by Japanese attendees at the last WG2 meeting,
> I quickly checked updates in WG2 N4982 (updated draft code
> table for Amd2 issued just before meeting) following IRG N2269
> (Editorial report of WS2015 IRG#50), and found so many
> problems.
>
> I report them in this message and hope it will help brushing up
> the code table. However please understand this is not a complete
> review due to the time constraint and the current draft is not
> matured yet. I'd like to suggest re-consider the schedule of CJK G
> progression.
>
>
>
> 2. Radicals
> #00593 u+3021E SC=8
> > The SC was counted 10 for unification of (争/爭8). Some other cases were processed similarly with references in IRGN221 which is continuously updated.
> #00748 u+30297 Ambiguous record. Should be deleted if no concensus.
> > The character is a variant of 葬. Current radical earth 土 was confirmed at IRG50 and in 2018-07, twice.
> #03538 u+30C19 Ambiguous record.
> > You are right. Radcial data should be changed to 166(里).
> #03852 u+30D31 Ambiguous record. 肉 or 言 ?
> > This was concluded 言 at IRG50.
>
>
> Below are not yet corrected as noted in the report.

> You are right. The comments of Zhuang chanracters (壯字) were confirmed after irg50 and their data should be modified but some of them (below) were neglected.
#00018
#00036
#00260
#00267
#00287
#00412
#00415
#00418
#00428
〉 #00428 was concluded radical fire 火 at IRG50 and was reflected in data. I also checked #00425 which is adjacent to #00428 in discussion record (IRGN2991 Appendix), its correction was reflected in data.
#00850
#00918
#00975
#01063
#01132
#01176
〉 #01175 was concluded radical meat 肉 at IRG50 and was reflected in data.
#01621
#01516
#01517
#01632
#01695
#01745
#01795
#01904
#01915
#02002
#02110
#02132
〉 #02132 was concluded radical water 水 at IRG50 and was reflected in data.
#02289
〉 #02289 was concluded radical fire 火 at IRG50 and was reflected in data.
#02351
〉 #02351 should be #02353.
#02355
#02681
#02711
#02823
〉 #02823 was concluded radical arrow 矢 at IRG50 and was reflected in data.
#03132

Inconsistent descriptions between "comment" and "note" which make it impossible to understand what is the conclusion.
#00434
〉 Mr. Eiso Chan (proxy of UK during IRG50) stated that the radical comments on #00434 and others were recommendations only, the the editorial meetings agreed to keep their current data.
#03582

〉 #03582 was concluded radical grass 艸 at IRG50 and was reflected in data. (The record "disagree" means to keep the current data.)

#04187

> #4187 was concluded radical cart 車 at IRG50 and was reflected in data.

#02087

> #2087 was concluded radical water 水 at IRG50 and was reflected in data. (The record "NO" means to keep the current data.)

#02038

> (The record "NO" means to keep the current data.)

#02039

> (The record "NO" means to keep the current data.)

#02393

> (The record "NO" means to keep the current data.)

#02437

> (The record "NO" means to keep the current data.)

#02849

> (The record "NO" means to keep the current data.)

#02997

> (The record "NO" means to keep the current data.)

#03009

> (The record "NO" means to keep the current data.)

#03334

> #03334 was concluded radical net 网 at IRG50 and was reflected in data.

4. Fontdesigns & normalizations

#00340 u+3010B SC=13

> You are right, sc=13.

#00623 u+30C01 Ambiguous record. Is it acceptable to change glyph ? If so, SC does not match to the glyphs.

> 呂 was counted 7 for unification to 吕. (The similar counting methods were applied to some exiting characters for unification possibilities in future.) It is a pity that we have to keep the G normailzed shape as 吕.

#00980 u+300D3 Right part is wrong.

> Let China recheck it.

#01023 u+3038A not yet corrected.

> The shape of 月 is correct. Ref.: G normalized 青.

#01174 u+3041C Wrong SC

> 羌 was counted 8 for unifications. Ref.: U+5534唴, U+7F8C羌, U+8723蜣, U+26351, U+263F1.

#01408 u+30504 Wrong SC

> 亟 was counted 9 for unifications. Ref: IRGN2221.

#02708 u+30940 This glyph is modified because of "PRC conventions". Same thing on u+30941(GZ-1982203) ?

> One of them should be modified to G normalized style. Let China recheck them.

#02868 u+309F4 Wrong SC

> Let China recheck its shape (􏰀石羌 or 􏰀石羌) and SC. (#02868 is at U+309CE in WG2N5006.)

#03008 u+30A35 Wrong SC

> You are right, sc=11.

#03140 u+30A9F It is questionable on "PRC conventions" to change 日 to 月

> Let China recheck it.

#03359 u+30B5F Glyph is not modified. Is it OK ? I wonder if there is no suggestion to modify the shape of 「及」 to match the evidence.

> Let Korea recheck it.

#03825 u+30D11 I wonder why the shape of「見」was simplified in WS2015 ver5.0. Is the current status correct ?

> The IDS was corrected from ⿴见吞 (Ver. 4.0) to ⿴見吞 (Ver. 5.0) as request of Eiso and UTC. Therefore, the radical T/S flag was consequently changed from R147.1=>R147.0, no more changes.

#04021 u+30DCA I guess 「喪→喪」 is a big change and it should not be "PRC conventions".Wrong SC as well.

> Let China recheck it.

#04111 u+30E21 It should not be PRC conventions on this change. Wrong SC.

> The shape is exactly the same as that in the dictionary, and follows the PRC conventions. It is a variant of 踆 according to the dictionary.

#04135 u+30E37 This change is beyond unifiable range and not a "normalization". Wrong SC as well.

> Let China recheck it.

7. Editorial Erros
7.1
#02972 u+3040B Radical is not changed to 爻

> You are right. (WS2015 SN is #2330.)

#03555 u+30C28 UK-01950 is not deleted.

> Please raise this in ballot. I guess that UK may ask move it WS2017 after balloting.

7.2
#00731 u+3027E Ambiguous record (About PDAM2.2 only) UK source is not added.

> Michel, please add UTC-01211. (This not UK, it is in the excel file.)

#00854 u+302F6 Ambiguous record (About PDAM2.2 only) UK source is not added.

> Michel, please add UK-01357. (This is in the excel file.)

7.4
#03461 u+30BD2 Source information is not modified as described in the record.

> Let China correct the source image or correct the data.

#04015 u+30DC5 Source information is not modified as described in the record.

> Let China correct the source image or correct the data.

#04032 u+30DD6 Source information is not modified as described in the record.

> Let China correct the source image or correct the data.

[General concerns]
Discussion record should be carefully reviewed before registration of it
to make the conclusions be clear. This document is not just a memo
but a part of Editorial report (Official document). Chinese words
such as 點 or 提 should be translated to English word.

At this time, China accepts so many "PRC conventions" or "normalization"
comment to modify glyphs. It looks very strange to me and I'd like
to suggest IRG re-consider the handling about "conventions" and
"normalization" issue on glyph design.

> Let China consider it.

I strongly request to setup the review period to end after "report is issued",

"updated glyphs are submitted" and "code table is updated". There
are so many errors every time in general.
> We may need one mor round review, but can we follow up the ballot schedule? Do we need to confirm
the ballot comments in IRG51 (Hanoi, Vietnam 2018-10-22/26)?

--
山本 知 （やまもと さとし）
gmail: yama.motch@gmail.com