

Disposition of the MSR1 comment made by TF-AIDN (Sarmad Hussain)

Last updated: May 7, 2014

Summary

This document describes a proposed disposition of the comments by TF-AIDN (Sarmad Hussain) posted on the ICANN web site as <http://forum.icann.org/lists/comments-msr-03mar14/msg00006.html>, as part of the Public Comment of the Maximal Starting Repertoire (MSR1) of the Root Zone Label Generation Rule Procedure (<http://www.icann.org/en/news/public-comment/msr-03mar14-en.htm>).

Each section shows the comments made by TF-AIDN along with the changes requested, and the disposition by the Integration Panel. The dispositions address only comments that request an action by the Integration Panel or a change to MSR1.

1 TF-AIDN Comment 2:

Based on its own experience, TF-AIDN re-iterates that an in-depth analysis of each code point being excluded by MSR1 is required by the relevant community. Therefore, TF-AIDN suggests that complete MSR1 should not be frozen together, but Integration Panel should consider freezing it one script at a time, after [each] relevant script Generation Panel has been formed and given its feedback on its relevant portion.

1.1 Requested Change

This effectively requests the removal from MSR1 of any script or repertoire not commented on by a Generation Panel

1.2 Disposition

Since this topic is of larger scope, the Integration Panel intends to respond separately.

2 TF-AIDN Comment 3

TF-AIDN requests IP to share reference materials they use internally for relevant discussions and decisions for MSR and LGR, to whatever extent possible. This can be done by posting the materials at the IP or IDN project web pages. Opening up this information would help understanding the IP perspective and for community and GPs to consult and incorporate such material timely. This will also help reach consensus and serve as a repository for future revisions of the MSR and LGR.

2.1 Requested Change

No change to the MSR repertoire or WLE rules.

2.2 Disposition

The Integration Panel has a policy of documenting source materials in the Rationale and Overview document. In case of inadvertent omissions, additional references and citations will be added.

3 TF-AIDN Comment 5

ICANN is also requested to work with IP to define a tentative timeline for releasing further versions of MSR so that language communities can organize their work accordingly, if revision in MSR1 is needed for a certain script.

3.1 Requested Change

No change to the MSR repertoire or WLE rules

3.2 Reply

The [Procedure] does not define the details of this mechanism. It is up to ICANN staff to define scheduling based on input from the various parties.

The Integration Panel will work with ICANN on this.

4 TF-AIDN Comment 6

TF-AIDN intends to look at Whole Label Rules toward the end of its Arabic Generation Panel work later this year. Thus, the task force has not reviewed the file MSR-1-Repertoire+WLERules-20140225.xml released along with MSR1 and reserves its comments on its contents at this time. We hope that IP will be able to incorporate feedback on this part of the work by TF-AIDN at a later stage.

4.1 Requested Change

No change to the MSR repertoire or WLE Rules

4.2 Disposition

The Integration Panel notes that the MSR1 XML document contains not only the default WLE rules, but constitutes the formal specification of the MSR repertoire. The PDF files, on the other hand are non-normative descriptions of the same repertoire derived algorithmically from the XML file. Once MSR1 is finalized, the XML file will only be updated as part of a future version of the entire MSR (MSR2). The *default* WLE rules for the MSR will also be finalized with the release of MSR1. However, each Generation Panel is entitled, under the Procedure, to propose additional WLE Rules relevant to its script as part of the LGR it proposes.

5 TF-AIDN Comment 7 – U+06AC ARABIC LETTER KAF WITH DOT ABOVE

TF-AIDN has the following specific comments on the Arabic script portion of the MSR1, in the context of its role as the Arabic Script Generation Panel.

The comment requests that U+ 06AC ARABIC LETTER KAF WITH DOT ABOVE be added to the MSR based on the fact that it is “used for Malay, Harari”. The comment further states: “use in Malay needs to be

Integration Panel: Disposition of the MSR1 comment made by TF-AIDN

further investigated. However, evidence of use is presented to consider the code point for MSR1.” The comment includes an image from a book and a keyboard layout and continues: *“However, the IDN table submitted by MyNIC in 2009 (see https://www.iana.org/domains/idntables/tables/my_ms-my_1.0.pdf) and the Malaysian Standard for Jawi (MS2443, 2012) does not include U+06AC but U+0762. Kew (2005) suggests that both variations are in use, though the latter should be preferred.”*

5.1 Proposed Change

Add U+06AC ARABIC LETTER KAF WITH DOT

5.2 Disposition

The comment indicates that the requested code point is an alternate for another code point U+0762, and, while attested, is not considered the preferred one. If this code point were to be included in the LGR it would create issues that would require resolution via the addition of a variant. The Integration Panel is more disposed towards keeping a less-preferred variant out of the LGR in the first place, than to contemplate creating a variant mapping.

The Integration Panel disposition is to reject the requested addition.

6 TF-AIDN Comment 6 – U+06E8 ARABIC SMALL HIGH NOON

The comment requests that 06E8 ARABIC SMALL HIGH NOON be added to the MSR1 and notes: *“Used in Chami:ni. Banafunzi and Vianello (2014) indicated that: “Nasal Sounds (at both the beginning of the words and the beginning of syllables) are noted above the letter following them (mb,nd), or altogether omitted in writing; The Chami:ni rhotarized group (nd/ndʻ) is reproduced with ﺝ with super imposed ﺏ (pg. 302).”*

6.1 Proposed Change

Add U+06E8 ARABIC SMALL HIGH NOON

6.2 Disposition

According to the source cited in the comment, the case for U+06E8 is rather weak, as *Chami:ni (or perhaps more properly Chimwiini (code: shw)) is a mostly oral dialect of Swahili. Banafunzi and Vianello write: “[it is] essentially an oral language. Except for private correspondence, and one example of Chimi:ni prose, writing of this language in Arabic script is confined to the Stenzi”(a form of poem).*

The Integration Panel disposition is to reject the requested addition.

7 TF-AIDN Comment 7– U+06F2 EXTENDED ARABIC-INDIC DIGIT TWO

The comment requests to include U+6F2 EASTERN ARABIC-INDIC DIGIT TWO in the MSR because it is *“used in Jawi for reduplication”* but that *“inclusion should be limited to only the case below, through Whole Label Rules and supporting documentation. General use of the code point should not be allowed.*

The comment further states:

Jawi uses the code point U+06F2 to indicate reduplication in the language. The plural is formed through this reduplication process, as are some other words. See <https://www.iana.org/domains/idn->

Integration Panel: Disposition of the MSR1 comment made by TF-AIDN

[tables/tables/my_ms-my_1.0.pdf](#) for some additional explanation. This is a productive process and no alternate way exists in representing this mechanism in Jawi.

The comment gives usage information according to the Malaysian Jawi Standard (MS2443, 2012; pg. 5), and continues: *As discussed in this standard, inclusion should be limited to only this case, through Whole Label Rules and supporting documentation. General use of the code point should not be allowed.*

7.1 Proposed Change

Add U+06F2 EASTERN ARABIC-INDIC DIGIT TWO to the MSR

7.2 Disposition

The general category of this characters is Nd, that is, the character is a digit. Under the Procedure, the Integration Panel is required to follow the Letter Principle. Digits are therefore not part of its mandate. Therefore these code points are not eligible to be added to MSR1.

The Integration Panel disposition is to reject the requested addition.

8 TF-AIDN Comment 7 – U+06FD-U+06FE Sindhi word symbols

The comment requests the addition of U+06FD ۞ ARABIC SIGN SINDHI AMPERSAND because it is “actively used in Sindhi to mean ‘and’” and of U+06FE ۞ ARABIC SIGN SINDHI POSTPOSITION MEN because it is “actively used in Sindhi to mean ‘in’ or ‘into’”.

The comment states further: *Sindhi uses ۞ for the conjunction “and” and ۞ for the post-position “in” or “into”. Both words are frequently used in Sindhi and no alternate spellings exist in writing them in Sindhi.*

As an example, introductory page of a recently published dictionary of Sindhi by Sindhi Language Authority is given below (Baloch 2006). Frequent use of both characters can be observed. (See comment for the relevant image)

8.1 Proposed Change

Add U+06FD ARABIC SIGN SINDHI AMPERSAND and U+06FE ARABIC SIGN SINDHI POSTPOSITION MEN to the MSR

8.2 Disposition

The requested characters have the Unicode General_Category of So, and are therefore classified as symbols by Unicode. Under the Procedure the Integration Panel is required to follow the Letter Principle, and symbols are not part of its mandate for the root (even if they are PVALID). They are further excluded from the set of code points recommended for identifiers in Unicode by their lack of the `XID_Continue` property.

Therefore, these code points are not eligible to be added to MSR1.

The Integration Panel disposition is to reject the requested addition.

Integration Panel Membership

This document is released by the Integration panel, composed of Asmus Freytag, Michel Suignard, Will Tan, Nicholas Ostler and Marc Blanchet.

References

[Procedure] Internet Corporation for Assigned Names and Numbers, "Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels." (Los Angeles, California: ICANN, March, 2013) <http://www.icann.org/en/resources/idn/variant-tlds/draft-lgr-procedure-20mar13-en.pdf>

[RFC5892] Faltstrom, P., Ed., "The Unicode Code Points and Internationalized Domain Names for Applications (IDNA)", RFC 5892, August 2010.

[Unicode63] The Unicode Consortium. The Unicode Standard, Version 6.3.0, defined by: "The Unicode Standard, Version 6.3.0", (Mountain View, CA: The Unicode Consortium, 2013. ISBN 978-1-936213-08-5). <http://www.unicode.org/versions/Unicode6.3.0/>.