Response of IP on Telugu LGR (Proposal of 2008-07-25)

DATE: 2018-08-02

# Overview

# All but two issues raised in previous IP response appear to have been incorporated. In the case of variants with Sinhala, the IP understands that the GP has a different position. We are repeating our feedback here for the record, but perhaps the best way to resolve this might be to present the proposal to the community for comment. In the case of the WLE rules, the IP notes that there is no disagreement between GP and IP on the effect of these rules, however, the rule for the Halant is unnecessarily complex in its XML implementation and the IP feels that it would be better to simplify this before submitting the LGR to public comment.

# Conclusion

# Excepting some unnecessary complexity in the XML implementation of the rules, the documents appear otherwise to be ready for public review.

There are two issues where the text of the document does need to be corrected, they are stated below and in a marked-up copy of the document.

There are a few minor copy edit issues that could be fixed as well (also provided in form of a marked up copy of the proposal document).

# Comments on main document (.docx)

|  |  |  |
| --- | --- | --- |
| **Item** | **Issue** | **IP Comment** |
| New-1 | **There are several minor editorial issues.**  See accompanying marked-up copy of the proposal document for a number of suggestions for improved wording. | Please review |
|  |  |  |
| New-2 | In section: “4.2 How to Avoid Duplicate Domain Names Involving ZWJ and ZWNJ?”  The sentence:  “In future if required, depending on the prevailing requirements by the community, the future NBGP may consider revisiting this rule. “  should be deleted. The exclusion of joiner characters is part of the Procedure for the RZ-LGR and not something that can be revisited. | Please review and fix |
|  |  |  |
| New-3 | In Section 7, the wording of Rule 5 no longer matches the implementation. It should be:  Rule 5. H cannot follow Nasal-C (Ref. Section 6.2 Type 1) | Please review and fix. |
|  |  |  |
| Old-3 | Cross-script variants with Sinhala: | The IP understands that the GP reviewed this and decided to make no change.  The issue is noted here for reference. |
|  |  | |
|  | The set above shows the only consonant that is proposed as a variant with Sinhala. In the opinion of the IP the similarity between the Telugu (0C30) and Sinhalese (0DBB) code points appears sufficiently remote to no longer be considered a homoglyph.  The same applies to the Kannada (0CB0) / Sinhala (0DBB) pair.  With only a single consonant (plus two combining marks) the overlap between Sinhala and these two scripts appears rather limited.  As was the case with similar relation between SEA scripts, the IP would recommend dropping Sinhala 0DBB from the variant sets for both Kannada and Telugu.  If that is done, the only remaining characters are combining marks 0D82 and 0D83 which can no longer form a cross-script label. They could then also be removed. | Please reconsider the case for 0DBB as a cross-script variant. It appears marginal to the IP. If GP agrees to remove, 0D82 and 0D83 would also no longer be necessary.  (Same feedback applies to Kannada). |

# Comments on LGR specification (.xml)

|  |  |  |
| --- | --- | --- |
| **Item** | **Issue** | **IP Comment** |
| Rules | If the suggested simplification of the rules for H is adopted, change rules in XML document to match.  The new rule would become  follows-C-except-Nasal-C  or in the regex notation  ([:C:]\[:Nasal-C:] ,  using a set difference between the character classes for C and Nasal-C. The rule would become a “when” instead of a “not-when” rule.  This would replace the compound rule “does-not-follow-C-or-follows-Nasal-C-and-precedes-C” and both its subsidiary rules “does-not-follow-C” and “follows-Nasal-C-and-precedes-C”. | The IP considers the current implementation, while formally correct in a logical and syntactic sense to be unnecessarily complex; this may make it impossible for a reviewer or implementer to understand the intent of the GP.  Please review again and fix.  (See the attached XML file for an implementation of the suggestion on the left.) |

# Comments on Test Labels

|  |  |  |
| --- | --- | --- |
| **Item** | **Issue** | **IP Comment** |
|  | The test labels have been reviewed. All valid labels report as valid, some invalid labels are resolved as invalid for different reasons than stated in the test file, but all invalid labels are reported as such.  One of the variant labels is reported as invalid. | Looks generally OK. |
|  | The IP evaluated the proposed simplified XML against the test file and got the same results (except, of course, for the name of the context rule where it occurs in the log file) |  |