To the chair

UNICODE TECHNICAL COMMITTEE

நூ த லோகசுந்தரம்

N.D.LOGASUNDARAM
selvindls61@gmail.com
14.4.2013

L2/13-065

The Sanskrit language must not be cloned inside the digital world

Once again I am here to reinstate my view on encoding of Grantham script of Sanskrit. It is in pipeline and records show as no action taken during last two meetings. The subject is in pending and suspense continued.

In the series of Unicore mail group postings on this Grantham that runs in many lengthy ones and the same subject were discussed by several members in detail, covering variety of the points which does not happen in the case of other encoding attempts.

In my previous L2 docs I have shown how the Sanskrit script of Grantham to be unified with DEVANAGARI block shown here again as a review.

Consortium by its own prime principles shall adopt Character encoding and not Glyph encoding. Even the presentation forms are also not entertained in the view that they have to be dealt in font (rendering) layer

Grantham merely being another font (Glyph set) of Sanskrit it is not eligible for new code points in any plane as those characters are already in DEVANAGARI block of BMP which is in stream and dynamism for a decade.

The present proposal of GoI being the case of request of duplicate codes in SMP for one and the same (congruent) Sanskrit characters which are already in DEVANAGARI blocks (main & extended) along with codes under Vedic tones.

There are views shown during discussions in UNICORE mail group postings on differences as leading and trailing positions of few vowel modifiers that are in the way of progress. It was already shown there itself as those are all due to hard scripted (defined) positional features provided under domain of DEVANAGARI. When DEVANAGARI is shared by Sanskrit language which happens to be with more than one glyptic forms with small differences in rendering (ortho) between them in few of its vowel modifiers, the hard scripted (defined) features infringes and pose as bug. Those fewer incompatibles are to be debugged by proper approach which will drive appropriate algorithmic solutions by the concerned technical faculties because nobody can have domain inside Sanskrit or any other existing language to change edit mutate native characters or modify any of its canonical properties or instigating any sort of damage direct or indirect to them in the process of putting it in the digital world. Always shoe is designed and devised for the foot.

Redefine the buggy orthographic positional (character) properties with bicameral perception to every character that in need (almost all the vowel modifiers) around 10 to 12 instead of the present hard scripted mono properties was already suggested.

There is no participation by GoI the final crystallized proposer connected in these intricate points in Unicode mail exchanges so far.

One subscriber seems to a font smith who is not for changes to DEVANAGARI character properties which are already defined and in use for his/her own interest to edit some scripting to suit new properties in the OS.

To the same subscribers view on a character it was shown without any doubt and replied in the same string of group mails that the eligibility of a character does not stem from a script whereas the language begets a character and forms parent / real owner / a title holder. A script is only a non-evaporative non-real-time format for eye and archiving. A language without a script is possible while a script without a language is a non entity. A language can be present in native

script as well as switched to another script in a transliteration mode which is a very simple mechanical equating process whereas the translation i.e. switching a language is a very complicated and multiplexing inelegant process connected to complex semantic layers

In any walks of life there will be inertia to change from existing facilities ambiences and modes they are to be driven with required force from manage mental tools. If there is a requirement anybody has to rise to the occasion and adopt. Because they feel difficult and react to insert some minor alterations = instead of finding the solutions try to escape or bye pass and ask for totally new path at somebody's expenses.

Next,
Even in the extreme case of impossibility of redefining with providing bicameral mode and it's real practicability by 100% demonstrated in the present scenario all the needed characters for play without the buggy properties can be allotted with dedicated code points anew was also suggested in the same group discussions

Similar dedicated characters were provided to cater different fonts as URUDU MARATHI SINDHI MARWARI KASHMERE etc for special conditions and needs within the shared DEVANAGARI range.

There are provisions made to suit Dravidian specials also as short E short O LLLA RRA NNNA with new created glyptic forms in rhythm with DEVANAGARI glyphs (with suitable combining ortho-graphic features inside DEVANAGARI) for same special functions. These will help some aspirant proposers as a fall out who want to write every Indian language contents in Grantham when it is unified with DEVANAGARI because those features were already taken care of by the presence in that DEVANAGARI. Even few objections seen in some docs claim them as Tamil characters to induce a bug inside Grantham like letters from GoTN and others become tangential because it is going to use the existing Sanskrit properties inside shared DEVANAGARI and also with entirely new different glyph forms nothing connected with Tamil.

In my doc I have shown rhythmic non-confusable glyphs for Grantham LLLA RRA NNNA for which I believe there cannot be any varied opinions even by GoTN etc which is very much as Grantham and not like Tamil form at all to confuse (as provided in proposals)

Regarding the ligating consonant cluster properties as in CV, CCV, CCCV the differing ortho properties are not defined in DEVANAGARI also as there are hundreds of permutation combination between combining consonants with their plurality in duality. Hense these variations are taken care of at the glyph rendering scripting and not connected with character coding layer. Those were dependant only by of vendors of different OS. Even in any stray case pops up again here again new dedicated character slot can be provided

Now,
Even in the event of not having sufficient vacant slots in DEVANAGARI range as we have allowed passing many of slots for encoding other initiatives in queue after the view of unification was presented still few code points from adjacent BMP zones can easily be provided for.

And even further,
if there is again any type hindrances crop up technical or otherwise, the surrogate pairs are always ready suitable and tangible with quantity and quality

Hence there are ways to adhere to the prime principle of character encoding avoiding at any cost and means for going in to glyph encoding which is a taboo.

Unicode is always stands for its own basics that if there is way and possibility to render a / set of character/s in the present setup no other alternates are entertained in allotting new character/s for complying the request whoever they may be whatever the case may be with what quantum of gravity behind them. They apply it without fail even in presentation forms cases of few characters in a range. Even here itself the proposer duly complied this obligation in their request for word final pure consonants with differing canonical presentation forms (by the presence of possibility of alternate rendering process they does not claim for separate code points)

In this case of GoI's Grantham proposal it is highly inappropriate in asking for a full-fledged and comprehensive full set of very long list of duplicate Sanskrit characters it was shown with a simile of someone trying to have a new spouse when there is one already very much alive and virile that

too with a condition for very long list of SREEDHANAA as found in one local group mail postings on Grantham in Unicode.

Going for a full fledged comprehensive duplicate character set afresh and away from what is already in DEVANAGARI for Sanskrit will be detrimental to that language itself. This will be nothing but an attempt to CLONE that language and paves way for BIRTH of

BMP Sanskrit # 1 & SMP Sanskrit# 2

In other words

Sanskrit of NORTH INDIA and SANSKRIT of Tamils

respectively and bring in a partisan attitude between the homogeneous commune on date in INDIA.

The handicaps imposed are not limited to the above.

Handling of Sanskrit contents will become costlier perennially in memory which cannot be brushed aside as it is from user point of view a vital and prime factor – their facilities (costing)

Huge volumes with long list of well archived titles in historical & orthodox devotional religious contents in 'MaNipravALam' a kind of texting mixing Sanskrit characters (in Grantham SCRIPT) into Tamil contents or vies-versa will not be possible.

(There are contents in much finer mixed mode also seen i.e. even inside a single word they are linguistically erratic as well as erroneous usage of characters that are not entertained by orthodox pedagogy of both the languages who with a goal of wisdom that is somewhat equal to colloquialy authored by a half baked cranky and irresponsible character/ glyph swappers for example writing

யு ni கோ du = Unicode in Tamil and English mixed mode

But in the cases of such fine tuned mixing mode used for impossible characters from their canonical repertoire of either of these languages for chosen unavoidable words are accepted condoned readily.

More nearer explanation for this point in focus the maNipravALam shall be the usage of French phrases words or Greek letters cum numerals & Arabic/Indian numerals inside English contents (a language of a content is decided by its syntax = sentence formation)

Applications in higher layers too than mere glyph rendering for which the Unicode codes should serve like voice-2-text, text-2-voice handling with logical operators as find and replace count N L P Machine translation etc are all will be very difficult to handle with bicameral duality in every element inside language contents start from naming and mention of the contents itself.

With all the above detailed and comprehensive submissions connected with the present proposal on Grantham encoding in pipe line by a the fulltime member GoI may request to resubmit back with an edited version to get the script into digital world quickly without any damage to Sanskrit without an iota of difficulty handicap etc in any way.

Thanking you, I remain

Chennai, INDIA   Yours Candidly

14.4.2013   N D LOGASUNDARAM

===

ADEPOT & ADROIT

Ingenious Engineers

===