Saurashtran Haaru consonants in Named Sequences list

1.0 Introduction

The speakers of Saurashtra language mostly live in the state of Tamil Nadu, India. In the majority usage, Tamil script with numbered diacritic marks is used to denote the Indo-Aryan aspirated and voiced consonants and print Saurashtran language texts. There is also a special Saurashtran script which is about a century old. Its standardiser was T. M. Rama Rao (1852-1913) who published some of his own books on Saurashtra grammar, a dictionary and song books in this distinctive script between 1899 to 1908 CE. Scholars like Edgar Thurston, R. N. Kuppusvami have recorded the invention of Saurashtra script by Sri. Rama Rao. The details are mentioned in the paper by H. N. Randle (An Indo-Aryan language of South India: Saurashtra-bhasha, BSOAS, 11, 1, (1943), pp. 104-121). This Saurashtran script is getting encoded in Unicode in a future version later than 5.0, currently in Draft Amendment 3 to ISO 10646.

2.0 Named Sequences of Saurashtra Haaru consonants

H. N. Randle (1943) describes a sub-letter 'upaaksharam' called 'haaru' to produce a particular kind of aspirated consonants: "Haarun - nh, mh, rh, lh occur in very few words, but most of them are words in frequent use: rhaa- (as common in Saurashtran as rah- in Hindi); mhoTTo (great); lhovvo (red); nhii- and nhaa- in negative forms of huu--; nhanno (small)."

In the WG document (N3059), the combining sign Haaru of the Saurashtra script is given a code point at U+A8B4. For the glyph shape of SAURASHTRA CONSONANT SIGN HAARU, refer to page 38, N3059 at
http://std.dkuug.dk/JTC1/SC2/WG2/docs/N3059.pdf

Two conjunct letters of Tamil, 'shrii' and 'kssa' conjuncts are getting added into Named Sequences in Unicode.
<U+0B95, U+0BCD, U+0BB7> TAMIL LETTER KSSA
<U+0BB6, U+0BCD, U+0BB0, U+0BC0> TAMIL LETTER SHRII
as given in the provisional list,
http://www.unicode.org/Public/5.0.0/ucd/NamedSequencesProv.txt
Following an approach similar to the mentioned Tamil letters, it is requested that the four Saurashtran Haaru consonants (Table 1) also be added into Named Sequences list.

Table 1. Named Sequences of Saurashtra Haaru Consonants

Number	Saurashtra Letter Name	Code Sequence with Haaru
1	SAURASHTRA LETTER NHA	<U+A8A5, U+A8B4>
2	SAURASHTRA LETTER MHA	<U+A8AA, U+A8B4>
3	SAURASHTRA LETTER RHA	<U+A8AC, U+A8B4>
4	SAURASHTRA LETTER LHA	<U+A8AD, U+A8B4>

3.0 Indic Named Consonant Sequences

The Saurashtra script has the regular aspirated consonant characters in its varga aksharams (kha, cha, Tha (retroflex), tha, pha) like any other Indo-Aryan language/script of India. In addition, it possesses a special class of aspirated consonants called Haaru (Reference 1) and these letters are generated in Unicode encoding with a Saurashtra consonant sign, Haaru (U+A8B4) which acts only a small subset of Saurashtran consonants. Because Saurashtra script is comparatively a recent script (References 1 and 2), if some more Haaru letters are discovered in print, coded sequences for those haarus can be added in Named Sequences, thus expanding Table 1.

The advantage of defining Haaru consonants in Named Sequences list is that unambiguous definition of attested haaru letters is made available for casual users, font developers, implementers and new students of the script.

In the future, named sequences for Latin letters with diacritic marks representing Indic letters in ISO 15919 have to be added (Reference 3). Another example of named sequences in Unicode is to make use of a MALAYALAM CONSONANT SIGN CILLU. Like Saurashtran combining sign haaru, Malayalam cillu sign works on a particular subset of consonants. Table 2 shows the Cillu glyph shapes of the nine base consonants attested Malayalam language texts and Table 3 shows Cillu consonants on 22 consonants required while printing Sanskrit texts in Malayalam script.

Table 2. Malayalam script Cillu consonants (page 92, Ref. 4)

k	ൿ & ൿ & ൔ	r	൱ & ൖ
ṇ	ൺ & ൺ	l	ൽ & ൽ
t	ൽ & ൽ & ൽ	ḷ	ൿ & ൿ
n	ൻ & ൻ	ḻ	ൔ
m	�മ		

Table 3. Tamil Grantha script Cillu consonants (page 16, Ref. 4)

k	...	n	...
g	...	p	...
ṅ	...	b	...
c	...	m	...
j	...	y	...
ñ	...	r	...
ṭ	...	v	...
ḍ	...	ś	...
ṇ	...	ṣ	...
t	...	s	...
d	...	h	...

If individual Cillus are to be retained in Unicode Malayalam code chart, canonical equivalences <Cillu consonant = Base consonant, Cillu sign> should be defined as being done for 11 letters in the Balinese script (Reference 5).

4.0 References

(1) Michael Everson, Correction of a Saurashtra character name in PDAM3, L2/06-091.
http://www.unicode.org/L2/L2006/06091-n3058-haaru.pdf
(2) Peri. Bhaskararao, Comments on Proposal to Encode the Saurashtra script in UCS, L2/03-277.
http://std.dkuug.dk/JTC1/sc2/WG2/docs/n2620.pdf
(3) Anshuman Pandey, Romanized Indic and LaTex, TUGboat, Vol. 19, 1998, No. 4.
http://www.tug.org/TUGboat/Articles/tb19-4/tb61pand.pdf
(4) Reinhold Grunendahl, South Indian Scripts in Sanskrit manuscripts and prints: Grantha Tamil - Malayalam - Telugu - Kannada - Nandinagari.
(2001: Harrassowitz Verlag, Wiesbaden, Germany).
(5) Balinese script code chart and decomposition of 11 of its characters,
http://www.unicode.org/Public/5.0.0/charts/CodeCharts-5.0.0d2.pdf