L2/03-067

Vedic Code Set

…a Draft

8.4 Unicode Standard for *Vedic Sanskrit* - a draft .....

Sanskrit has its own place recognized by the linguists all over the world.

India's ex-president and philosopher Dr Sarvapally Radhakrishan said, "Sanskrit as a language is an instrument of the greatest value in the delineation of all thought processes and the most profound rationisation of all ideas which are deep and subtle, of all forms of aesthetic and emotional perception, and above all, of the most profound and ultimate forms of intuition and understanding. It is agreed that the study of Sanskrit enables us to draw freely upon our tradition, which can lead to the new-world outlook of modern man. Further, it helps to keep pace with the rapid social change, advances in modern science & technology and the process of modernization, at the same time inculcating the right type of social, moral and spiritual value through self-discipline."

Sanskrit language is variously referred to as *Devavani, Amarvani, Girvanavani, Surbharati, Amarbharati*, etc. each expressing connoting its inherent vitality, versatility and greatness. The script in which Sanskrit texts are written is called Devanagari. It is claimed that the uniqueness of the Sanskrit language is that pronunciations of words, stanzas and sentences with measured intonation regulates ones giving and harmonizes once entire being with the subtle elements in the cosmic region.

If the flow of Sanskrit is arrested other languages, its branches, must perish for want of feeding. It is for this reason, perhaps, the great poet and *seear*, Ravindranath Tagore, desired that no professor should be in-charge of any language in Shantiniketan unless he is well grounded in Sanskrit.

Sanskrit is one of the most ancient languages of the world, which has molded the culture and the thought systems not only of India but also of many other countries in Asia. Sanskrit is not a dead language. Sanskrit was for over a millenium, a living spoken language with a considerable literature of its own. Besides works of literary value, there was a long philosophical and grammatical tradition. Sanskrit is still spoken in some Indian families. Even now new literature is being created in Sanskrit. Seventh system of philosophy, *Paramarthadarshan*, has been added to *Satdarshan* recently by Pundit Ramavatar Sharma. Its vocabulary has permeated all Indian languages, and thus provides continuity with the past of our country. There is renewed interest in learning Sanskrit because of its rich knowledge base in linguistics, philosophy, medicine, mathematics, astronomy, etc. Phonology (Study of Speech) and orthography (Study of Spelling) have not been so perfectly described in any natural language as in Sanskrit. Panini's book on Sanskrit Grammar, named Ashtadhyayi, has been considered by eminent American linguist Bloomfield as "one of the greatest monuments of human intelligence". Panini was preceded by a long chain of grammarians, and his tradition continued even afterwards. With his 4000 sutras, each of which is usually no more than two or three words, Panini was able to explain how almost all the words used in Sanskrit of his time were formed. It is precursor of today's generative grammar.

Sanskrit grammar is prescriptive, that takes phoneme as smallest unit with meaning, knowledge representation is deeper and holistic at sentence-level with three necessary and sufficient conditions of Expectancy (आकांक्षा), Compatibility (योग्यता), and Proximity (सन्निधि), whereas modern linguistics is descriptive and empirical, that takes morpheme as the smallest unit with meaning, uses word-by-word approach rather than sentence respectively.

Sanskrit is syntax-free and word-order-free natural language. Shastric Sanskrit is the Natural Language with all the desirable properties of formal artificial language, such as naturalness, expressiveness, un-ambiguity and no redundancy.

Karakas are important in case role assignment and thus facilitate in semantic extraction. There is equivalence between semantic net [representing tuples of verb, case role, and instance], and sentence analysis in Sanskrit. Knowledge inferencing in Sanskrit is therefore rather complete.

Knowledge is dealt with in *Apara Vidya Sastras* which are classified into four *Vedas* (scriptures), six *Vedanga-s* (Vedic auxiliary Science that deal with phonetics) and four *Upanga-s* (supplementary subjects).

Rig-Veda had 27 *Sakhas*, *Yajur Veda* had *Sukla: (15) & Krisna: (86) Sakhas*, *Sama Veda* had 1000 *Sakhas and Atharva Veda 9 Sakhas*. Every *Veda* has 4 types of texts; *Samhita, Brahmana, Aranyaka* and *Upanishad*. There is special Vedic grammar, rules for each *Sakha* known as *Prati Sakhya* and phonetic rules known as *Siksha*.

There are four *Upangas: Mimansa Sutra-s* (described rules for interpretation of *Vedic* text), *Nyaya & Vaisheshika sutra-s* (deal with logical aspects, ontological classification, process of human understanding), *Purana-s* (narrations of messages and teachings of Veda-s, *Dharma Sastra-s* (describe code of conduct for universal harmony).

There are 26 parameters for each *Vedic* syllabic definition.

Rick Briggs, a computer scientist of NASA in USA, published a paper in the Artificial Intelligence Magazine, 1985 on "Knowledge Representation in Sanskrit and Artificial Intelligence". He demonstrates that a natural language can serve as an artificial language such as Esperanto also, and that much work in Artificial Intelligence has been re-inventing what existed more than two thousand years ago. He establishes parallelism between modern scheme of knowledge representation using semantic nets and Sanskrit Grammarian's unambiguous sentence analysis. Modern knowledge-based computing employ Predicate Logic, Semantic Networks, Conceptual dependency schemes to represent World Knowledge. This may be related to *Sabda-bodha* concept dealt with in *Nyay, Vyakarna* and *Mimansa*.

Scientific & technological innovations which are contained in Sanskrit are given below in the following chronological table:

Period	S&T Innovations in Sanskrit
1500 B.C.	*Rigveda*: concept of natural law (rta): 1028 hymns & 10,462 richas
1000 B.C.	Samveda: book of melodies *Yajurveda*: the book of Sacrificial formulas. The whole series of 27 or 28 naksatras. Number names upto 10^{12} *Atharveda:* astronomical knowledge, more detailed medical Knowledge
1000 B.C.-500 B.C.	*Brahmanas*, *Aranyakas* and *Upanishads* doctrine of punchabhutas; Codification of medical knowledge into *Ayurveda* *Vedanga Jyotish*: 5 year cycle *Sulba-sutras:* beginning of geometry, irrational number Early ideas of *Vaisheshika*, *Samkhya* & *Mimamsa*; of *Bauddha*, *Jain* and *Charaka darshanas* Physical concepts: atomism, space, time, motion and sound

	Astronomical ideas: mathematical series (AP&GP)
	Agricultural practices to increase soil fertility
400 B.C.-400 A.D.	*Ayurvedic treatises* - Charaka and Sushruta Samhitas; Tridosha theory; extension of the doctrine of 5 elements, space, time and sound
	Arthashastra of Kautilya, Pingala's Chandah - sutra: Permutation, combinations and Binomial ideas
500 A.D.-1500 A.D.	*Nyaya Bhashya* of Vatsyayana: extension of atomic ideas, vision, sound, impetus theory; classification of animals and plants
	Padartha dharmasamgraha of Prashastapada: atomism, space, time, motion, sound
	Aryabhatta: theory of rotation of earth, epicycle theory of planetary motions, values of pie & sines, square & cube roots, indeterminate equation of the first order
	Panchasidhantika of Varahamihira
	Ganitasarasamgraha
	Amarakosa: classification and synonyms of plants and animals, minerals and metals
	Authoritative compilation of Ayurvedic knowledge; urine and pulse examination, Siddha system of medicine
	polytechnics: alchemical ideas; iron-casting, paper-making

Sanskrit Speech and Text

Sanskrit Grammar has distinguished the terms *varna* (phoneme) and *akshara* (syllable). Both these terms are used in the context of spoken languages and written languages respectively.

Since the oral tradition in India was of a higher order, the stress on right pronunciation was laid at most on the spoken language. To represent such speech nuances in written language, various *chinhas* (signs) evolved as to strike the equivalence in spoken and written expressions. This extra-ordinary activity is part of the Indian tradition. Therefore, the realization of such a system in the context of new technology seems to be imperative where writing is talked in the context of speech and speech in the context of writing. The attempt is made to identify *varnamala* comprising of basic speech sound units as vowel phonemes (*swara varna*) and consonant phonemes (*vyanjan varna*). These phonemes (*varnas*) when combined as C. ..C + V or only V form complete phonetic cluster. The correspondence in spoken and written syllables must be preserved through the Vedic Sanskrit Encoding scheme firstly by giving each phoneme a distinct code and secondly by giving each *chinha* -denoting nuances of speech -a distinct code.

Thus the scheme presented here comprises of following elements.

1. Phonemes -vowels, consonants
2. *Chinhas*
3. Punctuation marks
4. Digits

Sanskrit Phonology and Orthography

The Devanagari script is used for writing classical Sanskrit as well as Vedic Sanskrit. This includes the multi-tier usage of diacritic marks of complex compositions, above, below and at the sides of the base glyphs. Therefore, as compared to modern

historical derivatives from Sanskrit such as *Hindi, Marathi, Nepali* etc., the Sanskrit text demands adequate range of characters as well as exhaustive rendering rules to achieve the advanced typographic quality in Vedic Sanskrit text. The provision of additional codes from U+0800 -U+08FF is provided to address these issues.

Encoding principles : The effective unit of the Sanskrit writing system is the phoneme (*varna*). The range of phonemes (*Varnamala*) consists of '*SwaraVarna*' (Vowel Phoneme) and '*Vyanjan Varna*' (Consonant phoneme). While '*Swara Varna*' is self-powered and it is not dependent on any other element, the '*Vyanjan Varna*' however, needs an addition of '*Swara Varna*' to compose a syllabic entity. While '*Swara Varna*' (V) can be written down as syllables ('*akhara*'), other syllables ('*akharas*') are the outcome of the combination of '*Vyanjan Varna*' and '*Vowel Varna*'.

Phoneme (*varana*) to Syllables (*aksharas*)

As mentioned earlier phonemes are divided into two types: vowel phonemes (*swara varna*) and consonant phonemes (*vyanjan varna*). They together broadly constitute the *Varnamala* which has been referred as a *varna-samamnaya*. The orthographic representation of these *varnas* is done in a systematic way. The combination of consonant phoneme and a vowel phoneme produces a syllable (*akshara*). A cluster of glyphs emerges as an outcome of this process.

For example,

/k/ + /a/ = /ka/ syllable which is written as ... क्+अ= क

/p/ + /aa/ = /paa/ syllabic *akshara* is /paa/ प्+आ= पा

Please note that

Corresponding to each *swara* phoneme there is an *akshara* which is its syllabic form.

Vowel phoneme अ आ इ ई

Vowel syllable अ आ इ ई

Rendering of *aksharas* (Syllables)

k-phoneme + /a/ = k-*akshar* क्+अ=क

The syllables formed by adding vowel phonemes /a/, /aa/, /i/, etc. to the consonant phoneme are written by creating *aksharas*. One consonant phoneme added to all the *swara* phonemes one by one is called a *baaraakhadi*.

Thus the concept of extended range of 'Barakhadi' (12 syllables) is achieved in the following way.

K(d) + vv1 = K + A = KA	क्+अ=क
K(d) + vv2 = K + AA = KAA	क्+आ=का
K(d) + vv3 = K + I = KI	क्+इ=कि
K(d) + vv4 = K + II = KII	क्+ई=की
K(d) + vv5 = K + U = KU	क्+उ=कु
K(d) + vv6 = K + UU = KUU	क्+ऊ=कू
K(d) + vv7 = K + Vocalic R = K(Vocalic)R	क्+ऋ=कृ
K(d) + vv8 = K + Vocalic RR = K(Vocalic)RR	क्+ॠ=कॄ
K(d) + vv9 = K + Vocalic L = K(Vocalic)L	क्+ ऌ=कॢ
K(d) + vv10 = K + Vocalic L = K(Vocalic)LL	क्+ॡ=कॣ
K(d) + vv11 = K + E = KE (Short)	क्+ऐ=कॅ
K(d) +vv12 = K + EE = KE	क्+ए=के
K(d) + vv13 = K + E = K(Candra}E	क्+ऍ=कॅ
K(d) + vv14 = K + AE = KAI	क्+ऐ=कै
K(d) + vv15 = K + O = KO (Short)	क्+ ऑ=कॊ
K(d) + vv16 = K + O = KO	क्+ओ=को
K(d) + vv17 = K + O = K(Candra)O	क्+ऑ=कॉ
K(d) + vv18= K + AU=KAU	क्+औ=कौ

Syllables can also be formed by adding vowel phonemes to a sequence of more than one consonant phonemes. These syllables are called jodaksharas or sanyuktaksharas. For example :

k-phoneme + y-phoneme + aa-phoneme = kyaa
क् + य् + आ = क्या

s-phoneme + t-phoneme + u-phoneme = stu
स् + त् + उ = स्तु

Please note that the invariant element in this process is the set of phonemes. The variation occurs in the shape of glyphs written in various Indian scripts. For example, the phoneme /k/ and /0/ will result in the glyph shape where graphic element is added in front and on the top where as in Bengali, graphic shape will be added in front and prior to the base glyph. Therefore this model can be extended to most of the Indian languages which have phonetic base. To sum up the proposed scheme calls for code points for consonant phoneme *k* as compared to the existing Devanagari code which provides code points for glyph *ka*. The proposed scheme is of additive nature (k + a) as compared to subtractive model. This scheme would allow unambiguous representation of the entire repertoire of characters required in creating the exhaustive Devanagari script syllabic range along with its phonetic values.

Vedic Vagvarna Chart

The split up of the codes in chart 1 and chart 2. The Chart 1 would facilitate *Laukik Sanskrit* while the remaining symbols are included in Chart 2.

Chart 1. The *Laukik Sanskrit Varnas* and *Chinhas*

It is possible to achieve the following using the Chart 1.

1. Text composition in the Indian languages that use Devanagari script
2. Transliteration (Indirect method)
3. Transliteration of other Indian languages into Devanagari script.
4. Sort and search in *Laukik Sanskrit*

Chart 2: The other *Vedic Sanskrit Varnas* and *Chinhas*

The Chart 2 will facilitate (along with chart 1) composition of text from all four Vedas with their intonation Vedic marks and the phonetic break-up of words from dictionaries

The two charts presented here are designed taking into consideration following character coding requirements:

1. Non-ambiguity
2. Transliteration
3. Phonetic break-up
4. Exhaustiveness
5. Uniqueness
6. Backward compatibility
7. Default sorting for lankik sanskrit

Code Set Design Considerations

1. The range of S*anskrit Vedic* marks, which were identified after extensive research have been included.
2. All consonant letter signs have been shown as pure consonants (characters with *HALANT* refered elsewhere as dead consonants).
3. Devanagari script in its excluded form as mentioned in *Manak Hindi Vartani* (Standardised Hindi) issued by Central Hindi Directorate has been taken as reference.
4. Adequate *Bhedak Chinhas* are provided to take care of phonetic variations of *Kashmiri, Urdu, Sindhi, South Indian Languages, Persian* and *Arabic.*
5. In addition to the *Bhedak Chinhas,* reserved space is provided in the code chart for incorporation of different phonemes in future.
6. Vowel letters such as A, AA, I etc. shown in the chart have been included for their phonetic character (value). The respective vowel *matras* are not explicitly represented as they can be derived unambiguously through positional logic (CV, CCV, etc.).
7. The IPA equivalence for these Indian Phonetic Letter signs can be established.
8. The range of *Swaraadi-Anuswaar* and *Visarga* used for *Laukik Sanskrit*, are kept in the 1st chart. The total range of *Anuswaar* and *Visarga*

as needed for Vedic Sanskrit text have been placed separately in the 2nd chart.

9. The total marks in terms of *Udaatt, Anudaatt, Swarit and Swarit Kamp* have been put next to S*waraadis*. From the total range of five *Kamp* signs namely, *Hrasva Kamp, Dirgha Kamp, Udaatt Kamp, Tathabhavya Kamp* and *Shiva Kamp* as mentioned in '*Shiksha*' first two have been given code points. The remaining three codes have been reserved for three remaining *Kampas*.

10. The *Samvedic Swarochar* signs have been incorporated representing different schools of *Samavedic* traditions. This scheme would facilitate to compose *Samvedic* intonation marks on the top of the syllable through a program.

11. The codes from 08F0 to 08FF have been reserved for Speech Control Commands in context of Text to Speech and Speech to Text technology.

12. This coding scheme would facilitate to create total repertoire needed to compose *Vedic* text *(Rigved, Yajurved, Atharvaved* and *Samaved)*.

13. The sort order for *Vedic* text, if needed has to be handled by using specific algorithms.

Lexical order and sorting

Following decisions were taken to facilitate sorting in chart 1.

1. The numerals have been put in the beginning to suit the international convention.

2. The necessary additional signs (pitch, stress, time) for phonetic breakup usage have been put in the page 2 of the chart.

3. The logical order of vowel phonemes followed by vowel phoneme modifiers and consonant phonemes followed by consonant phonemes modifiers has been followed. The *Anuswaar* and the *Visarg* group is kept after vowel phonemes modifiers. This would ensure correct sorting in *Laukik (Abhijaat)* -Classical Sanskrit.

4. Through the chart, transliteration from other Indian languages to Devanagari is possible through indirect mapping methodology.

5. Phonetic break up approach has been taken for the words in dictionary. No sort order in phonetic break up of words is needed.

6. *Poorna Viram Chinha* is differentiated from the *Dashamsha Chinha and* Viyog *Chinha* is differentiated with the *Sanyog Chinha*.

7. The *Runa* Chinha and *Gunaka* Chinha are provided.

The Salient Features of Vedic Vagvarna Encoding Scheme

- The new scheme of Phonemes (vowels and pure consonants) as character codes for Unicode, is nearer to the linguistic model and serves all the linguistic needs.

- There is a provision to extend this to newer combinations not yet identified (tribal and folk languages).

- The text-processing operations like indexing and sorting which are very important for information storage and retrieval on computers can be performed efficiently.

- Speech synthesis can be facilitated as the nuances of speech are preserved through these encoding.

- An absolute requirement on any script encoding is that it be possible for a computer to take any valid sequence of underlying character codes and algorithmically render the appropriate visual form, given a repertoire of surface glyphs. In the case of *Vedic Sanskrit* encoding scheme presented here the required character shaping rules are well-formed and therefore font rendering systems can been built based on this.

Vedic Code Chart 1

	080	081	082	083	084	085	086	087
0	० 0800	R 0810	ऐ 0820	R 0830	R 0840	द् 0850	R 0860	◌ 0870
1	ऀ 0801	इ 0811	R 0821	क् 0831	झ् 0841	ध् 0851	R 0861	ARLU 0871
2	२ 0802	R 0812	ओ 0822	R 0832	R 0842	न् 0852	ल् 0862	◌़ 0872
3	३ 0803	ई 0813	ओ 0823	ख् 0833	ञ् 0843	R 0853	R 0863	◌़ 0873
4	४ 0804	उ 0814	R 0824	R 0834	ट् 0844	R 0854	व् 0864	◌़ 0874
5	५ 0805	R 0815	ऑ 0825	ग् 0835	ठ् 0845	प् 0855	R 0865	◌़ 0875
6	६ 0806	ऊ 0816	औ 0826	R 0836	ड् 0846	फ् 0856	श् 0866	० 0876
7	७ 0807	R 0817	R 0827	घ् 0837	R 0847	R 0857	ष् 0867	ऽ 0877
8	८ 0808	ऋ 0818	ॅ 0828	ङ् 0838	R 0848	ब् 0858	स् 0868	। 0878
9	९ 0809	ॠ 0819	◌ॆ 0829	च् 0839	ढ् 0849	R 0859	ह् 0869	SNDH 0879
A	अ 080A	ऌ 081A	ॆ 082A	R 083A	R 084A	भ् 085A	ळ् 086A	ARSN 087A
B	R 080B	ॡ 081B	ॉ 082B	छ् 083B	ण् 084B	म् 085B	R 086B	RIKT 087B
C	R 080C	ऐ 081C	ॅ 082C	R 083C	R 084C	य् 085C	R 086C	DRSK 087C
D	R 080D	ए 081D	ॉ 082D	ज् 083D	त् 084D	R 085D	व्ह 086D	R 087D
E	आ 080E	ऍ 081E	ः 082E	R 083E	R 084E	र् 085E	R 086E	R 087E
F	R 080F	R 081F	R 082F	R 083F	थ् 084F	R 085F	R 086F	R 087F

Vedic Code Chart 2

	088	089	08A	08B	08C	08D	08E	08F
0	0880	0890	08A0	08B0	08C0	08D0	08E0	08F0
1	0881	0891	08A1	08B1	08C1	08D1	08E1	08F1
2	0882	0892	08A2	08B2	08C2	08D2	08E2	08F2
3	0883	0893	08A3	08B3	08C3	08D3	08E3	08F3
4	0884	0894	08A4	08B4	08C4	08D4	08E4	08F4
5	0885	0895	08A5	08B5	08C5	08D5	08E5	08F5
6	0886	0896	08A6	08B6	08C6	08D6	08E6	08F6
7	0887	0897	08A7	08B7	08C7	08D7	08E7	08F7
8	0888	0898	08A8	08B8	08C8	08D8	08E8	08F8
9	0889	0899	08A9	08B9	08C9	08D9	08E9	08F9
A	088A	089A	08AA	08BA	08CA	08DA	08EA	08FA
B	088B	089B	08AB	08BB	08CB	08DB	08EB	08FB
C	088C	089C	08AC	08BC	08CC	08DC	08EC	08FC
D	088D	089D	08AD	08BD	08CD	08DD	08ED	08FD
E	088E	089E	08AE	08BE	08CE	08DE	08EE	08FE
F	088F	089F	08AF	08BF	08CF	08DF	08EF	08FF

Vedic Code Details

Sanskrit Devanagari Vedic Anka

0800	०	VEDIC SANSKRIT ANKA SHUNYA
0801	१	VEDIC SANSKRIT ANKA EKAN
0802	२	VEDIC SANSKRIT ANKA DVI
0803	३	VEDIC SANSKRIT ANKA TRI
0804	४	VEDIC SANSKRIT ANKA CHATUR
0805	५	VEDIC SANSKRIT ANKA PANCHAN
0806	६	VEDIC SANSKRIT ANKA SHASH
0807	७	VEDIC SANSKRIT ANKA SAPTAN
0808	८	VEDIC SANSKRIT ANKA ASHTAN
0809	९	VEDIC SANSKRIT ANKA NAVAN

Sanskrit Devanagari Swara Varna

080A	अ	VEDIC SANSKRIT SWARA VARNA A Unrounded central half-open
080B	R	Reserved
080C	R	Reserved
080D	R	Reserved
080E	आ	VEDIC SANSKRIT SWARA VARNA AA Unrounded central open
080F	R	Reserved
0810	R	Reserved
0811	इ	VEDIC SANSKRIT SWARA VARNA I Unrounded front close short
0812	R	Reserved
0813	ई	VEDIC SANSKRIT SWARA VARNA II Unrounded front close long
0814	उ	VEDIC SANSKRIT SWARA VARNA U Rounded back close short
0815	R	Reserved
0816	ऊ	VEDIC SANSKRIT SWARA VARNA UU Rounded back close long
0817	R	Reserved
0818	ऋ	VEDIC SANSKRIT SWARA VARNA VOCALIC R Consonant in the form of vowel short
0819	ॠ	VEDIC SANSKRIT SWARA VARNA VOCALIC RR Consonant in the form of vowel long
081A	ऌ	VEDIC SANSKRIT SWARA VARNA VOCALIC L Consonant in the form of vowel short
081B	ॡ	VEDIC SANSKRIT SWARA VARNA VOCALIC LL Consonant in the form of vowel long
081C	ऐ	VEDIC SANSKRIT SWARA VARNA E SHORT Unrounded front half-close long

081D	ए	VEDIC SANSKRIT SWARA VARNA E Unrounded front half-close long
081E	ऍ	VEDIC SANSKRIT SWARA VARNA E WITH CHANDRAKAR ABOVE Unrounded front half-open
081F	R	Reserved
0820	ऐ	VEDIC SANSKRIT SWARA VARNA AI Compound vowel (A + I)
0821	R	Reserved
0822	ओ	VEDIC SANSKRIT SWARA VARNA O SHORT Rounded back half-close long
0823	ओ	VEDIC SANSKRIT SWARA VARNA O Rounded back half-close long
0824	R	Reserved
0825	ऑ	VEDIC SANSKRIT SWARA VARNA AA WITH CHANDRAKAR ABOVE Rounded back half-open
0826	औ	VEDIC SANSKRIT SWARA VARNA AU Compound vowel (A + U)
0827	R	Reserved

Sanskrit Devanagari Swara Bhedak Chinha

0828	◌॑	VEDIC SANSKRIT SWARA BHEDAK CHNHA 1 for Kashmiri
0829	◌॒	VEDIC SANSKRIT SWARA BHEDAK CHINHA 2 for Urdu
082A	◌॓	VEDIC SANSKRIT SWARA BHEDAK CHINHA 3 for Avesta
082B	◌॔	VEDIC SANSKRIT SWARA BHEDAK CHINHA 4 for Avesta

Sanskrit Devanagari Swaraadi Chinha

082C	◌ँ	VEDIC SANSKRIT SWARADI CHINHA – CHANDRABINDU
082D	◌ं	VEDIC SANSKRIT SWARADI CHINHA – ANUSVARA
082E	◌ः	VEDIC SANSKRIT SWARADI CHINHA – VISARGA
082F	R	Reserved
0830	R	Reserved

Sanskrit Devanagari Vyanjan Varna

0831	क्	VEDIC SANSKRIT VYANJANA VARNA K Voiceless unaspirated velar plosive stop
0832	R	Reserved
0833	ख्	VEDIC SANSKRIT VYANJANA VARNA KH Voiceless aspirated velar plosive stop
0834	R	Reserved
0835	ग्	VEDIC SANSKRIT VYANJANA VARNA G Voiced unaspirated velar plosive stop
0836	R	Reserved

0837	घ्	VEDIC SANSKRIT VYANJANA VARNA GH Voiced aspirated velar plosive stop
0838	ङ्	VEDIC SANSKRIT VYANJANA VARNA NG Nasal velar plosive stop
0839	च्	VEDIC SANSKRIT VYANJANA VARNA C Voiceless unaspirated palatal plosive stop
083A	R	Reserved
083B	छ्	VEDIC SANSKRIT VYANJANA VARNA CH Voiceless aspirated palatal plosive stop
083C	R	Reserved
083D	ज्	VEDIC SANSKRIT VYANJANA VARNA J Voiced unaspirated palatal plosive stop
083E	R	Reserved
083F	R	Reserved
0840	R	Reserved
0841	झ्	VEDIC SANSKRIT VYANJANA VARNA JH Voiced aspirated palatal plosive stop
0842	R	Reserved
0843	ञ्	VEDIC SANSKRIT VYANJANA VARNA NY Nasal palatal plosive stop
0844	ट्	VEDIC SANSKRIT VYANJANA VARNA TT Voiceless unaspirated retroflex plosive stop
0845	ठ्	VEDIC SANSKRIT VYANJANA VARNA TTH Voiceless aspirated retroflex plosive stop
0846	ड्	VEDIC SANSKRIT VYANJANA VARNA DD Voiced unaspirated retroflex plosive stop
0847	R	Reserved
0848	R	Reserved
0849	ढ्	VEDIC SANSKRIT VYANJANA VARNA DDH Voiced unaspirated retroflex plosive stop
084A	R	Reserved
084B	ण्	VEDIC SANSKRIT VYANJANA VARNA NN Nasal retroflex plosive stop
084C	R	Reserved
084D	त्	VEDIC SANSKRIT VYANJANA VARNA T Voiceless unaspirated dental plosive stop
084E	R	Reserved
084F	थ्	VEDIC SANSKRIT VYANJANA VARNA TH Voiceless aspirated dental plosive stop
0850	द्	VEDIC SANSKRIT VYANJANA VARNA D Voiced unaspirated dental plosive stop
0851	ध्	VEDIC SANSKRIT VYANJANA VARNA DH Voiced aspirated dental plosive stop

0852	न्	VEDIC SANSKRIT VYANJANA VARNA N Nasal dental plosive stop
0853	R	Reserved
0854	R	Reserved
0855	प्	VEDIC SANSKRIT VYANJANA VARNA P Voiceless unaspirated bilabial plosive stop
0856	फ्	VEDIC SANSKRIT VYANJANA VARNA PH Voiceless aspirated bilabial plosive stop
0857	R	Reserved
0858	ब्	VEDIC SANSKRIT VYANJANA VARNA B Voiced unaspirated bilabial plosive stop
0859	R	Reserved
085A	भ्	VEDIC SANSKRIT VYANJANA VARNA BH Voiced aspirated bilabial plosive stop
085B	म्	VEDIC SANSKRIT VYANJANA VARNA M Nasal bilabial plosive stop
085C	य्	VEDIC SANSKRIT VYANJANA VARNA Y Voiced palatal semi-vowel – vocalised consonant
085D	R	Reserved
085E	र्	VEDIC SANSKRIT VYANJANA VARNA R Voiced unaspirated alveolar flapped semi-vowel
085F	R	Reserved
0860	R	Reserved
0861	R	Reserved
0862	ल्	VEDIC SANSKRIT VYANJANA VARNA L Voiced dental lateral semi-vowel
0863	R	Reserved
0864	व्	VEDIC SANSKRIT VYANJANA VARNA V Voiced labio-dental semi-vowel
0865	R	Reserved
0866	श्	VEDIC SANSKRIT VYANJANA VARNA SH Voiceless palatal fricative
0867	ष्	VEDIC SANSKRIT VYANJANA VARNA SHH Voiceless retroflex fricative
0868	स्	VEDIC SANSKRIT VYANJANA VARNA S Voiceless alveolar fricative
0869	ह्	VEDIC SANSKRIT VYANJANA VARNA H Voiced glottal fricative
086A	ळ्	VEDIC SANSKRIT VYANJANA VARNA LL Voiced dental lateral semi-vowel
086B	R	Reserved
086C	R	Reserved
086D	ळ्ह	VEDIC SANSKRIT VYANJANA VARNA LLH
086E	R	Reserved
086F	R	Reserved
0870	◌	VEDIC SANSKRIT LUPTAK CHINHA

0871	ARLU	VEDIC SANSKRIT ARDH LUPTAK CHINHA

Sanskrit Devanagari Vyanjan Bhedak Chinha

0872	◌़	VEDIC SANSKRIT VYANJAN BHEDAK CHINHA 1 (ONE NUKTA) for Urdu, Malayalam, Tamil, Marathi
0873	◌़़	VEDIC SANSKRIT VYANJAN BHEDAK CHINHA 2 (TWO NUKTA) for Malayalam
0874	◌़़़	VEDIC SANSKRIT VYANJAN BHEDAK CHINHA 3 (THREE NUKTA) for Kashmiri, Avesta
0875	◌़	VEDIC SANSKRIT VYANJAN BHEDAK CHINHA 4 (LINE WITH ONE NUKTA) for Sindhi
0876	०	VEDIC SANSKRIT CHINHA – SANKSHEPA
0877	ऽ	VEDIC SANSKRIT SWARADI CHINHA – AVAGRAHA
0878	।	VEDIC SANSKRIT CHINHA – DANDA
0879	SNDH	VEDIC SANSKRIT CHINHA – SAANDHAKA
087A	ARSN	VEDIC SANSKRIT CHINHA – ASAANDHAKA
087B	RIKT	VEDIC SANSKRIT CHINHA – RIKTAKA
087C	R	Reserved
087D	DRSK	VEDIC SANSKRIT CHINHA – DARSHAK
087E	R	Reserved
087F	R	Reserved

Vedic Sanskrit Phonetic Break-up Signs

0880	◌॑	VEDIC SANSKRIT UCCHA SWAN
0881	◌॒	VEDIC SANSKRIT NIMNA SWAN
0882	◌॓	VEDIC SANSKRIT BALA (STRESS) BHEDAK CHINHA 1
0883	◌॔	VEDIC SANSKRIT BALA (STRESS) BHEDAK CHINHA 2
0884	◌ऀ	VEDIC SANSKRIT KAAL (TIME) BHEDAK CHINHA ATI LAGHU
0885	◌ँ	VEDIC SANSKRIT KAAL (TIME) BHEDAK CHINHA LAGHU
080D	◌ं	VEDIC SANSKRIT KAAL (TIME) BHEDAK CHINHA GURU
0887	◌ः	VEDIC SANSKRIT SWARADI CHINHA – CHANDRKOR WITH ARDHA ANUSWAR Partial nasalization indicator (soft)
0888	◌ऺ	VEDIC SANSKRIT SWARADI CHINHA – ARDHA ANUSWAR Partial nasalization stress indicator

Vedic Sanskrit Anuswar

0889	◌ऻ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC (KRISHNA) ANUSVAR 1
088A	◌़	VEDIC SANSKRIT ANUSVAR –YAJURVEDIC (KRISHNA) LONG ANUSVAR 2

088B	◌ॆ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR MADHYA 3
088C	◌ॆ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR DAKSHIN 4
088D	◌ॆ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 5
088E	◌ॆ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC (SHUKLA) ANUSVAR 6
088F	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 7
0890	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 8
0891	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 9
0892	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 10
0893	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 11
0894	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 12
0895	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 13
0896	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 14
0897	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC ANUSVAR 15
0898	◌ॅ	VEDIC SANSKRIT ANUSVAR – ANUSHMANS ANUSVAR 16
0899	◌ॅ	VEDIC SANSKRIT ANUSVAR – YAJURVEDIC (SHUKLA) ANUSVAR 17
089A	R	Reserved

Vedic Sanskrit Swaraadi Chinha – Visarga

089B	◌ः	VEDIC SANSKRIT ARDH VISARGA
089C	◌ः	VEDIC SANSKRIT VISARGA
089D	÷	VEDIC SANSKRIT VISARGA 1
089E	◌ः	VEDIC SANSKRIT VISARGA 2
089F	◌ः	VEDIC SANSKRIT VISARGA 3
08A0	◌ः	VEDIC SANSKRIT VISARGA 4
08A1	◌ः	VEDIC SANSKRIT VISARGA 5
08A2	◌ः	VEDIC SANSKRIT VISARGA 6
08A3	◌ः	VEDIC SANSKRIT VISARGA 7
08A4	◌ः	VEDIC SANSKRIT VISARGA 8
08A5	◌ः	VEDIC SANSKRIT VISARGA 9

08A6	᳦	VEDIC SANSKRIT ARDHA-VISARGA – JIHVA-MULIYA 1
08A7	᳧	VEDIC SANSKRIT ARDHA-VISARGA – JIHVA-MULIYA 2
08A8	᳨	VEDIC SANSKRIT ARDHA-VISARGA – JIHVA-MULIYA 3
08A9	ᳩ	VEDIC SANSKRIT ARDHA-VISARGA – UPADHAMANIYA 1
08AA	ᳪ	VEDIC SANSKRIT ARDHA-VISARGA – UPADHAMANIYA 2
08AB	ᳫ	VEDIC SANSKRIT ARDHA-VISARGA – UPADHAMANIYA 3
08AC	R	Reserved
08AD	R	Reserved

Vedic Sanskrit TAARATA Chinha

08AE	◌᳭	VEDIC SANSKRIT UDATTA
08AF	◌ᳮ	VEDIC SANSKRIT ANUDATTA

Vedic Sanskrit Rigvedic Svarita Chinha

08B0	◌ᳯ	VEDIC SANSKRIT SVARITA - ATHARVA-VEDIC SVARITA
08B1	◌ᳰ	VEDIC SANSKRIT SVARITA UDATTA
08B2	◌ᳱ	VEDIC SANSKRIT SVARITA -LONG SVARITA
08B3	◌ᳲ	VEDIC SANSKRIT SVARITA – MAITRAYANI
08B4	◌ᳳ	VEDIC SANSKRIT SVARITA ANUDATTA
08B5	◌᳴	VEDIC SANSKRIT SVARITA - SHUKLA YAJURVEDIC SVARITA
08B6	◌ᳵ	VEDIC SANSKRIT SVARITA – MAITRAYANI SVARITA ONE
08B7	◌ᳶ	VEDIC SANSKRIT SVARITA – YAJURVEDIC SVARITA
08B8	◌᳷	VEDIC SANSKRIT SVARITA – KATTHAK/ MAITRAYANI SAMHITA JATYA SVARITA
08B9	◌᳸	VEDIC SANSKRIT SVARITA – MAITRAYANI SVARITA TWO
08BA	◌᳹	VEDIC SANSKRIT SVARITA - ANTIM UDATTA
08BB	◌ᳺ	VEDIC SANSKRIT SVARITA - KATTHAK ANUDATTA
08BC	◌ᳺ᳹	VEDIC SANSKRIT SVARITA – WITH 2-S SHAPES BELOW
08BD	ह᳽	VEDIC SANSKRIT SVARITA DEVENAGARI H WITH HORIZONTAL CROSSLINE (comes in Maitrayani Samhita jatya)
08BE	R	Reserved
08BF	R	Reserved

Vedic Sanskrit Svarita Kamp

08C0	◌᳾	VEDIC SANSKRIT SVARITA RHASVA KAMP

08C1	ॅ	VEDIC SANSKRIT SVARITA DEERGH KAMP
08C2	R	Reserved
08C3	R	Reserved
08C4	R	Reserved
08C5	R	Reserved
08C6	R	Reserved

Vedic Sanskrit Samavedic Swarochchar Chinha

08C7	ं	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA ONE
08C8	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA TWO
08C9	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA THREE
08CA	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA FOUR
08CB	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA FIVE
08CC	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA SIX
08CD	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA SEVEN
08CE	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA EIGHT
08CF	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA NINE
08D0	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA KAMPA
08D1	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA R
08D2	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA ONE R
08D3	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA TWO R
08D4	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA THREE R
08D5	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA FOUR R
08D6	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA FIVE R
08D7	ः	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA R KAMPA

08D8	◌॒	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 1R KAMPA
08D9	◌॓	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 2R KAMPA
08DA	◌॔	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 3R KAMPA
08DB	◌॒॒	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 4R KAMPA
08DC	◌॓॓	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 5R KAMPA
08DD	◌॑	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA U
08DE	◌॒॑	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 2U
08DF	◌॒॑	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA K
08E0	◌॑॓	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA 3K
08E1	◌॑॔	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA URDHVA
08E2	◌॒॒॑	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA SMALL AVAGRAHA
08E3	◌॓॑॓	VEDIC SANSKRIT SAMAVEDIC SWAROCHCHAR-CHINHA stress
08E4	R	Reserved

Vedic Sanskrit Special Symbols

08E5	ॐ	VEDIC SANSKRIT SYMBOL OM
08E6	॥	VEDIC SANSKRIT SYMBOL SWASTIK
08E7	◌॒॒॑॓॔	VEDIC SANSKRIT CHINHA - PUSHPIKA
08E8	◌॒॒॒॒॑॓॔॑॓॔	VEDIC SANSKRIT VEDIIC CHINHA – AVAGRAHA-DWAYA
08E9	॥	VEDIC SANSKRIT CHINHA – DANDA DWAYA
08EA	½SP	VEDIC SANSKRIT CHINHA – KHANDA (HALF SPACE)
08EB	...	VEDIC SANSKRIT LOPA CHINHA
08EC	◌॒॒॒॒॒॒॑॓॔॑॓॔॑॓॔	VEDIC SANSKRIT KAAKAPADA
08ED	-॥-	VEDIC SANSKRIT TATHAIVA
08EE	×	VEDIC SANSKRIT GUNAKA CHINHA
08EF	·	VEDIC SANSKRIT DASHANSH CHINHA
08F0	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F1	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F2	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F3	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)

08F4	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F5	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F6	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F7	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F8	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08F9	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08FA	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08FB	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08FC	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08FD	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08FE	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)
08FF	R	Reserved for Vaka Niyaman Sammadesh (Speech Control Commands)

SANSKRIT-DESHA Keyboard 2002
DESHA-Multilingual Phonemic Keyboard

Designed and Developed by National Centre for Software Technology, Mumbai, India
Concept and Layout : Prof. R.K. Joshi (Revised version of Vividha 1986, Desha 1990)
Other language specific keyboards may follow the layout with suitable modifications.

(Courtesy : Prof. R. K. Joshi, Visiting Design Specialist at NCST, Juhu, Mumbai-400049, India.
Email: rkjoshi@konark.ncst.ernet.in.
Tel: 26201606, 6201574. Fax: +91-22 6210139.
in consultation with
Dr. Sadanandan, Vice President, Bharati Sanskrit Vidya Niketanam, Ghatkopar, Mumbai
Dr Alka Irani, Sr Research Scientist, NCST and other Sanskrit Scholars.)