BrahmiNet-ITRANS Transliteration Scheme

Anoop Kunchukuttan

August 2020

The BrahmiNet-ITRANS notation (Kunchukuttan et al., 2015) provides a scheme for transcription of major Indian scripts in Roman, using ASCII characters. It extends the ITRANS[1] transliteration scheme to cover characters not covered in the original scheme. Tables 1a shows the ITRANS mappings for vowels and diacritics (*matras*). Table 1b shows the ITRANS mappings for consonants. These tables also show the Unicode offset for each character. By Unicode offset, we mean the offset of the character in the Unicode range assigned to the script. For Indic scripts, logically equivalent characters are assigned the same offset in their respective Unicode codepoint ranges. For illustration, we also show the Devanagari characters corresponding to the transliteration.

ITRANS	Unicode Offset	Devanagari
a	05	अ
aa, A	06, 3E	आ, ा
i	07, 3F	इ, ि
ii, I	08, 40	ई, ी
u	09, 41	उ, ु
uu, U	0A, 42	ऊ, ू
RRi, R^i	0B, 43	ऋ, ृ
RRI, R^I	60, 44	ॠ, ॄ
LLi, L^i	0C, 62	ऌ, ॢ
LLI, L^I	61, 63	ॡ, ॣ
.e	0E, 46	ऎ, ॅ
e	0F, 47	ए, े
ai	10,48	ऐ, ै
.o	12, 4A	ऒ, ॊ
o	13, 4B	ओ, ो
au	14, 4C	औ, ौ
aM	05 02, 02	अं
aH	05 03, 03	अः
.m	02	ं
.h	03	ः

(a) Vowels

ka 15 क	kha 16 ख	ga 17 ग	gha 18 घ	~Na, N^a 19 ङ
cha 1A च	Cha 1B छ	ja 1C ज	jha 1D झ	~na, JNa 1E ञ
Ta 1F ट	Tha 20 ठ	Da 21 ड	Dha 22 ढ	Na 23 ण
ta 24 त	tha 25 थ	da 26 द	dha 27 ध	na 28 न
pa 2A प	pha 2B फ	ba 2C ब	bha 2D भ	ma 2E म
ya 2F य	ra 30 र	la 32 ल	va, wa 35 व	
sha 36 श	Sha 37 ष	sa 38 स	ha 39 ह	
Ra 31 ऱ	lda, La 33 ळ	zha 34 ऴ		

(b) Consonants

Version: v1.0 (9 August 2020)

NOTES:

1. The ITRANS scheme has the same character for the independent vowel and the corresponding vowel diacritic (*matra*).
2. For consonants, the Roman character represents the consonant without the implicit *schwa* vowel. For instance, *k* represents क्. To represent the implicit vowel, *a* has to follow the consonant (*ka* represents क).

[1] https://en.wikipedia.org/wiki/ITRANS

CITATION: If you are using the BrahmiNet-ITRANS scheme, please cite the following work:

```
@inproceedings{kunchukuttan2015brahmi,
  title = "{Brahmi-Net: {A} transliteration and script conversion system
      for languages of the Indian subcontinent}",
  author = "Anoop Kunchukuttan and Ratish Puduppully and Pushpak Bhattacharyya",
  booktitle = "NAACL: System Demonstrations",
  year = "2015",
}
```

API: You can generate BrahmiNet-ITRANS transliteration using the *Indic NLP Library*[2] (Kunchukuttan, 2020)

References

Kunchukuttan, A. (2020). The IndicNLP Library. https://github.com/anoopkunchukuttan/indic_nlp_library/blob/master/docs/indicnlp.pdf.

Kunchukuttan, A., Puduppully, R., and Bhattacharyya, P. (2015). Brahmi-Net: A transliteration and script conversion system for languages of the Indian subcontinent. In *NAACL: System Demonstrations*.

[2] https://github.com/anoopkunchukuttan/indic_nlp_library