Draft Policy Document
For

INTERNATIONALIZED DOMAIN NAMES

Language: MANIPURI

Department of Information Technology,
Ministry of Communications and Information Technology,
Government of India, New Delhi

Contents

1. AUGMENTED BACKUS-NAUR FORMALISM (ABNF)

1.1 Declaration of Variables:

Dash   → Hyphen -

Digit   → Indo-Arabic digits [0-9]

C   → Consonant

V   → Vowel

M   → Matra

D   → Anusvara

B   → Chandrabindu

X   → Visarga

H   → Halant

N   → Nukta

Z   → Khanda ta

1.2 ABNF Operators

Sr. No.	Operator	Function
1	"/"	Alternative
2	"[ ]"	Optional
3	"*"	Variable Repetition
4	"( )"	Sequence Group

In what follows, the Vowel Sequence and the Consonant Sequence pertinent to Manipuri are given. To facilitate understanding, equivalents in Devanagari are provided.

1.3 Vowel Sequence

A vowel sequence is made up of a single vowel. It may be followed but not

necessarily (optionally) by an Anuswara (D),Candrabindu (B) or a Visarga (X) The number of D, B or X which can follow a V in MANIPURI may not be restricted to one.The possibility of a Visarga or Anusvara following a Chandrabindu does not exist in MANIPURI.

The vowel sequence in MANIPURI is therefore V [D |B |X]

Examples :

V	অ	अ
V[D]	অং	अं
V[B]	অঁ	अँ
V[X]	অঃ	अः

1.4   Consonant Sequence

A consonant sequence admits the following shapes:
1. A single consonant (C)
Example:

C	क	क

2. A consonant optionally followed by dependent vowel sign[M] or anuswara[D] or Chandrabindu[B] or visarga[X] or Halanta [H]

C[M|D|B|X|H]

Example:

C[M]	কি	कि	
C[D]	কং	कं	
C[B]	কঁ	कँ	
C[X]	কঃ	कः	
C[H]	ক্	क्	(Pure Consonant)

2.a. A CM sequence can be optionally followed by D, B or X
(CM)[D | B | X]

Example:

CM[D]	कीं	कीं
CM[B]	काँ	काँ
CM[X]	वीः	वीः

3. A sequence of consonants (up to 3) joined by Halant (also known as Hasant/Virama).

*2(CH)C
Example:

CHCHC   त्र्य = त् + र् + य   त्+र्+य

Subsets
3.a. The combination may be followed by M, D, B or X
 Example:

CHC[M]	क्की	क ् क ी	क्की	क ् क ी
CHC[D]	क्कं	क ् क ं	क्कं	क ् क ं
CHC[B]	क्कँ	क ् क ँ	क्कँ	क ् क ँ
CHC[X]	क्कः	क ् क ः	क्कः	क ् क ः

3.b. *2(CH)CM may be followed by D, B, or X
*2(CHCM)[D | B | X]
Example:

CHCM[D]	क्कीं	क ् क ी ं	क्कीं	क ् क ी ं
CHCM[B]	क्काँ	क ् क ा ँ	क्काँ	क ् क ा ँ
CHCM[X]	क्कीः	क ् क ी ः	क्कीः	क ् क ी ः

4. A single Khanda Ta(Z)
Example:

Z   ৎ

The final canonical structure of the consonant sequence in IDN can be defined in ABNF as: *2(C[N]H)C[N][H|D|B|X|M[D|B|X]] | Z

1.5   ABNF Applied to the MANIPURI IDN

The formalism can be applied to create/validate IDN labels. So a valid IDN label can be defined as follows.

Vowel-sequence → V [D |B |X]
Consonant-sequence → *2(C[N]H)C[N][H|D|B|X|M[D|B|X]] | Z
Sequence → consonant-sequence | vowel-sequence
IDN-label → (sequence | digit) * ([dash] (sequence |digit))

Additional Examples putting more light on MANIPURI ABNF:

Below are some of the examples which will help a casual reader understand some of the rules ABNF puts in place. These are just given for reference purposes and are not meant to be comprehensive
1.   H |M |D|B|X cannot occur in the beginning of an IDN domain name
Example:

ْক   ়ক
িক   িক
ংক   ংক
ঁক   ঁক
ঃক   ঃক

As can be seen they will result automatically in a "golu" marking an invalid character. This is an intrinsic property of the Indic syllable and is quasi automatically applied wherever supported by the OS.

2. H is not permitted after V, D, B, X, M, digit and dash
    Example

अ्   अ्
कं््   कं्

कँ◌ॕ   कँ◌ॕ

कः◌ॕ   कः◌ॕ

क्ि   क्ि

1◌ॕ   1◌ॕ

−◌ॕ   -◌ॕ

3. Number of D,B or X permitted after consonant-sequence or vowel-sequence or M is restricted to one.

   Example

कं◌ॅ   कं◌ं

कँ◌ँ   कँ◌ँ

कः◌ः   कः◌ः

कीं◌ॅ   कीं◌ं

काँ◌ँ   काँ◌ँ

कीः◌ः   कीः◌ः

अं◌ॅ   अं◌ं

अँ◌ँ   अँ◌ँ

अः◌ः   अः◌ः

4. Number of M permitted after consonant-sequence is restricted to one
Example:

की◌ी   की◌ी

5. M is not permitted after V
Example:

औा   ईा

6. The combination of Anuswara+Visarga as well as Visarga+Anuswara is not permissible
Example:

कं◌ः   कं◌ः

कः◌ॅ   कः◌ं

2. RESTRICTION RULES

The Augmented Backus Naur Formalism (ABNF) is generic in nature and when applied to a specific language/script, certain restriction rules apply. In other words, in a given language some of the Formalism structures do not necessarily apply. To take care of such cases restriction rules are set in place. These restrictions will help to fine-tune the ABNF.

In the case of MANIPURI the following rules apply:

1. Khanda ta is NOT allowed in the beginning of an IDN label.

2. Nukta can be only allowed after following characters:
    ড (09A1)
    ঢ (09A2)
    য (09AF)

3. A consonant sequence that is intended to end with Halant [H] can only be followed by Hyphen or Digit. Thus following combinations are permissible.

    ক্-
    ক্1

4. Consecutive Hyphens will not be permitted in a domain name.

5. The number of consecutive identical consonants joined by a Halant within a label shall not exceed two. Thus ক্ক (ka+halant+ka) is permitted but not ক্ক্ক (ka+halant+ka+halant+ka).

6. A label containing not more than three "akshara", which have got variants shall be permitted. As an example let us consider a, b, c and d as four aksharas in a given label having a', b', c' and d' as variants in which case such a label will be disallowed. (Example of disallowed label - abcd, acdb,

cdaba and so on)

Additional Note:

Wherever a variant is present in a given label, the variants shall be strictly symmetric and non-transitive. Thus given a variant ৗী and ৗীঁ, the variants of a label such as রনী shall be রনীঁ. রনীঁঁ generated by adding an extra ৗীঁ to ৗী shall not be permitted regardless of its ABNF Validity status. This ensures that over generativity does not take place.

3. EXAMPLES

Following are the some examples to show MANIPURI IDN Label.

C	ক	কল
CN	ড়	গুড়
CH	প্	কপ্
CM	মী	মরী
CD	লং	লং
CB	কঁ	কঁ
CX	কঃ	কঃ
CMD	চাং	চাং
CMB	কাঁ	কাঁ
CMX	দিঃ	দিঃ
CHC	ক্ক	ক্ক
V	অ	অ
VD	অং	অং
VB	অঁ	অঁ
VX	অঃ	অঃ
Z	ৎ	ৎ

4. LANGUAGE TABLE: MANIPURI

0980   Bengali   09FF

Manipuri

5. NOMENCLATURAL DESCRIPTION TABLE OF MANIPURI LANGUAGE TABLE

CANDRABINDU (B)		
0981	ঁ	BENGALI SIGN CANDRABINDU
ANUSVARA (D)		
0982	ং	BENGALI SIGN ANUSVARA
VISARGA (X)		
0983	ঃ	BENGALI SIGN VISARGA
INDEPENDENT VOWELS (V)		
0985	অ	BENGALI LETTER A
0986	আ	BENGALI LETTER AA
0987	ই	BENGALI LETTER I
0988	ঈ	BENGALI LETTER II
0989	উ	BENGALI LETTER U
098A	ঊ	BENGALI LETTER UU
098F	এ	BENGALI LETTER E
0990	ঐ	BENGALI LETTER AI
0993	ও	BENGALI LETTER O
0994	ঔ	BENGALI LETTER AU
098C	ঌ	BENGALI LETTER L
098B	ঋ	BENGALI LETTER VOCALIC R
CONSONANTS (C)		
0995	ক	BENGALI LETTER KA
0996	খ	BENGALI LETTER KHA
0997	গ	BENGALI LETTER GA
0998	ঘ	BENGALI LETTER GHA
0999	ঙ	BENGALI LETTER NGA
099A	চ	BENGALI LETTER CA
099B	ছ	BENGALI LETTER CHA
099C	জ	BENGALI LETTER JA

099D	ঝ	BENGALI LETTER JHA
099E	ঞ	BENGALI LETTER NYA
099F	ট	BENGALI LETTER TTA
09A0	ঠ	BENGALI LETTER TTHA
09A1	ড	BENGALI LETTER DDA
09A2	ঢ	BENGALI LETTER DDHA
09A3	ণ	BENGALI LETTER NNA
09A4	ত	BENGALI LETTER TA
09A5	থ	BENGALI LETTER THA
09A6	দ	BENGALI LETTER DA
09A7	ধ	BENGALI LETTER DHA
09A8	ন	BENGALI LETTER NA
09AA	প	BENGALI LETTER PA
09AB	ফ	BENGALI LETTER PHA
09AC	ব	BENGALI LETTER BA
09AD	ভ	BENGALI LETTER BHA
09AE	ম	BENGALI LETTER MA
09AF	য	BENGALI LETTER YA
09B0	র	BENGALI LETTER RA
09B2	ল	BENGALI LETTER LA
09B6	শ	BENGALI LETTER SHA
09B7	ষ	BENGALI LETTER SSA
09B8	স	BENGALI LETTER SA
09B9	হ	BENGALI LETTER HA
09F1	ৱ	BENGALI LETTER RA WITH LOWER DIAGONAL
VOWEL SIGNS(MATRAS) (M)		
09BE	া	BENGALI VOWEL SIGN AA
09BF	ি	BENGALI VOWEL SIGN I
09C0	ী	BENGALI VOWEL SIGN II

09C1	্ু	BENGALI VOWEL SIGN U
09C2	ূ	BENGALI VOWEL SIGN UU
09C3	ৃ	BENGALI VOWEL SIGN VOCALIC R
09C7	ে	BENGALI VOWEL SIGN E
09C8	ৈ	BENGALI VOWEL SIGN AI
09CB	ো	BENGALI VOWEL SIGN O
09CC	ৌ	BENGALI VOWEL SIGN AU
09E2	ৢ	BENGALI VOWEL SIGN VOCALIC L

VIRAMA (H)

09CD	্	BENGALI SIGN VIRAMA

KHANDATA (Z)

09CE	ৎ	BENGALI LETTER KHANDATA

NUKTA (N)

09BC	়ঁ	BENGALI SIGN NUKTA

6. VARIANT TABLE FOR MANIPURI

Actual Character / characters	Variant Character / characters
ৃ (09C3)	ৢ (09C2)
ী ঁ (09C0) + (0981)	ী (09C0)
ৌ ঁ (09CC) + (0981)	ৌ (09CC)

7. EXPERTS CONSULTED/TO BE CONSULTED

Dr. Hidam Dolen, North Eastern Regional Language Centre

8. PROPOSED ccTLD FOR MANIPURI

India (Bhārat) localized in Manipuri - ভারত

Note : You can send your feedbacks to idn-feedback@cdac.in