Proposal to encode CHAKMA LETTER LHAA, DEPENDENT VOWEL SIGNS AA & EI for Chakma

Bivuti Chakma and Andrew Glass
(bsereye at hotmail dot com | anglass at outlook dot com)

November 2, 2016

Chakma script was added to Unicode in version 6.1 (2012) on the basis of Michael Everson’s proposal documents submitted in 2008 and 2009 (L2/08-133 & L2/09-187R). The present encoding for Chakma lacks three characters that are used by the Baarah Maaatraa orthography for Chakma language. The present document addresses the gap in the Chakma encoding by proposing these signs as additions to the existing Chakma block (U+11100–U+1114F). The three characters are:

<table>
<thead>
<tr>
<th>Glyph</th>
<th>Code</th>
<th>Character name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11144</td>
<td>CHAKMA LETTER LHAA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>= fetfada lhaa</td>
</tr>
<tr>
<td></td>
<td>11145</td>
<td>CHAKMA VOWEL SIGN AA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>= deil-dya aa</td>
</tr>
<tr>
<td></td>
<td>11146</td>
<td>CHAKMA VOWEL SIGN EI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>= dvi adi gadei-dya ei</td>
</tr>
</tbody>
</table>

Michael Everson’s original proposal was aware of the development of orthographies for Tanchangya language and anticipated the need for encoding additional characters. These signs are not being proposed in connection with Tachangya language. This document is not a statement regarding the completeness of the Chakma encoding for any other language or orthography other than Baarah Maaatraa orthography for Chakma language. The characters proposed here, Chakma LETTER LHAA, VOWEL SIGNS AA & EI have also been used traditionally in regional Chakma orthographies and are not simply modern innovations.

The LETTER LHAA is used for the aspirated counterpart to LETTER LAA (figures 1, 2, 3 & 4). This sign is a CV akṣara that may combine with dependent vowel signs, for example:



It can also occur in conjunct consonant clusters, for example:

 (lhyaa)

The VOWEL SIGN AA  is a dependent vowel sign used to mark an explicit long A vowel (see figure 4, 5, 6, 7, 8, 12, 13, 14 & 15) in some orthographies. It is related to the Myanmar sign TALL AA (U+102B). There is no corresponding independent form of this vowel.

The VOWEL SIGN EI  is a dependent vowel sign used to mark an explicit long E vowel mix with I vowel (see figure 9, 10, 11, 12, 13, 14 & 15). It is derived from the CHAKMA LETTER E (U+11106) & I (U+11104). There is no corresponding independent form of this vowel.
Character Data

Character properties
In the format of UnicodeData.txt:

```
11144;CHAKMA LETTER LHAA;Lo;O;L;;;;;N;;;;;
11145;CHAKMA VOWEL SIGN AA; Mc;O;L;;;;;N;;;;;
11146;CHAKMA VOWEL SIGN EI; Mc;O;L;;;;;N;;;;;
```

Linebreaking
In the format of LineBreak.txt:

```
11144;AL   # Lo    [1] CHAKMA LETTER LHAA
11145..11146;CM # Mc     [2] CHAKMA VOWEL SIGN AA..CHAKMA VOWEL SIGN EI
```

Syllabic categories
In the format of IndicSyllabicCategory.txt:

```
# Indic_Syllabic_Category=Consonant
11144; Consonant # Lo [1] CHAKMA LETTER LHAA

# Indic_Syllabic_Category=Vowel_Dependent
11145..11146; Vowel_Dependent # Mc [2] CHAKMA VOWEL SIGN AA..CHAKMA VOWEL SIGN EI
```

Positional categories
In the format of IndicPositionalCategory.txt:

```
# Indic_Positional_Category=Right
11145; Right # Mc   CHAKMA VOWEL SIGN AA

# Indic_Positional_Category=Right
11146; Right # Mc   CHAKMA VOWEL SIGN EI
```

Sorting
Sort values should be assigned to these characters so that the following sort order is achieved:

- \( \odot \) (U+11126) < \( \odot \) (U+11144) < \( \odot \) (U+11103)
- \( \odot \) (U+11134) < \( \odot \) (U+11145) < \( \odot \) (U+11146)
- \( \odot \) (U+11145) < \( \odot \) (U+11146) < \( \odot \) (U+11100)

References

Add the three additional characters:

1. Chakma Consonant letter LHAA
   = fetfada lhaa
   = U+11144

2. Dependent Vowel Sign AA
   = deil-dya aa
   = U+11145

3. Dependent Vowel Sign EI
   = dvi adi gadei-dya ei
   = U+11146
Various signs

11100 𑄦𑄳𑄣 CHAKMA SIGN CANDRABINDU
= caaraphuda

11101 𑄣 CHAKMA SIGN ANUSYARA
= ekaphuda

11102 𑄣 CHAKMA SIGN VISARGA
= dpiphuda

Independent vowels

11103 𑄣 CHAKMA LETTER AA
= phiphuda aa

11104 𑄣 CHAKMA LETTER I
= delabhaanipadaa i

11105 𑄣 CHAKMA LETTER U
= buhuma uu

11106 𑄣 CHAKMA LETTER E
= lepaahu e

Consonants

11107 𑄣 CHAKMA LETTER KAA
= cuuyangya kaa

11108 𑄣 CHAKMA LETTER KHA
= ghuuyangya khaa

11109 𑄣 CHAKMA LETTER GAA
= caandyaa gaa

1110A 𑄣 CHAKMA LETTER GHA
= maadlialyaa ghaa

1110B 𑄣 CHAKMA LETTER NGAA
= dllaama ngaa

1110C 𑄣 CHAKMA LETTER CAA
= chilbahaaya caa

1110D 𑄣 CHAKMA LETTER CHAA
= majana choa

1110E 𑄣 CHAKMA LETTER JAA
= dhyuyadlaa jaa

1110F 𑄣 CHAKMA LETTER JHAA
= uvaatraa jhaa

11110 𑄣 CHAKMA LETTER NYAA
= dbaayaa nyaa

11111 𑄣 CHAKMA LETTER TI AA
= dnyaadlaa taa

11112 𑄣 CHAKMA LETTER TIIA
= phudaadlyuait tiaa

11113 𑄣 CHAKMA LETTER DAA
= aadadaangst dhaa

11114 𑄣 CHAKMA LETTER DOAA
= lebharat dhaa

11115 𑄣 CHAKMA LETTER NAA
= petthiyaayaa nya

11116 𑄣 CHAKMA LETTER TAA
= gphongdaait taa

11117 𑄣 CHAKMA LETTER THAA
= japaadat thaa

11118 𑄣 CHAKMA LETTER DAA
= dolarit dhaa

11119 𑄣 CHAKMA LETTER DHAA
= talaamuayt dhaa

1111A 𑄣 CHAKMA LETTER NAA
= phiaraahnayaa nya

1111B 𑄣 CHAKMA LETTER PAA
= dphayaa paa

1111C 𑄣 CHAKMA LETTER PHAA
= ubaphuda phaai

1111D 𑄣 CHAKMA LETTER BAA
= ubamnuvya baa

1111E 𑄣 CHAKMA LETTER BHAA
= ciaoalayya bhaa

1111F 𑄣 CHAKMA LETTER MA
= bhuupdalaas ma

11120 𑄣 CHAKMA LETTER YYAA
= dmyuyya yya

11121 𑄣 CHAKMA LETTER YA
= jhuj yaa

11122 𑄣 CHAKMA LETTER RAA
= chilasaya raa

11123 𑄣 CHAKMA LETTER LA
= talanuyya laa

11124 𑄣 CHAKMA LETTER WAA
= buhonyaa waa

11125 𑄣 CHAKMA LETTER SAA
= buhilbukyaa saa

11126 𑄣 CHAKMA LETTER HAA
= ubamnuyya haa

Dependent vowel signs

11127 𑄣 CHAKMA VOWEL SIGN A
= ubutaliykaa a

11128 𑄣 CHAKMA VOWEL SIGN I
= baabhaat i

11129 𑄣 CHAKMA VOWEL SIGN II
= baaniyphudaat ii

1112A 𑄣 CHAKMA VOWEL SIGN U
= rukhumaat uu

1112B 𑄣 CHAKMA VOWEL SIGN UU
= duttanu uu

1112C 𑄣 CHAKMA VOWEL SIGN E
= okaraat e

1112D 𑄣 CHAKMA VOWEL SIGN AI
= delabhaangai ai

1112E 𑄣 CHAKMA VOWEL SIGN O
= okaraat o

1112F 𑄣 CHAKMA VOWEL SIGN AU
= aukaatau au

1112G 𑄣 CHAKMA VOWEL SIGN OI
= okaraat o

11130 𑄣 CHAKMA O MARK
= deel-yaa

11132 𑄣 CHAKMA AU MARK
= dvi adi gadei-yaa eii

Various signs

11133 𑄣 CHAKMA VIRAMA
• used to form conjuncts
  → 1039 𑄣 myanmar sign virama

11134 𑄣 CHAKMA MAAYYA
• killer
  → 103A 𑄣 myanmar sign asat

Digits

11136 𑄣 CHAKMA DIGIT ZERO

11137 𑄣 CHAKMA DIGIT ONE

11138 𑄣 CHAKMA DIGIT TWO

11139 𑄣 CHAKMA DIGIT THREE

1113A 𑄣 CHAKMA DIGIT FOUR

1113B 𑄣 CHAKMA DIGIT FIVE

1113C 𑄣 CHAKMA DIGIT SIX

1113D 𑄣 CHAKMA DIGIT SEVEN

1113E 𑄣 CHAKMA DIGIT EIGHT

1113F 𑄣 CHAKMA DIGIT NINE

11144 𑄣 CHAKMA LETTER LHAA
= lefadaa laa

11145 𑄣 CHAKMA VOWEL SIGN AA
= deli-yaa aa

11146 𑄣 CHAKMA VOWEL SIGN EI
= dvi adi gadei-yaa ei
Figure 1: Usage of the proposed character consonant letter **LHAA** of Chakma [from *Chakma Legha Shigheni Bazar 2011* (*Chakma Script learning year-2011*). Source: Online Published from Tripura, India]
Figure 2: Usage of the proposed character consonant letter LHAA of Chakma [from Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011). Source: Online Published from Tripura, India]
Figure 3: Usage of the proposed character consonant letter LHAA of Chakma [from 30 Minide Chakma Legha Shigho (Learn Chakma Script within 30 minutes). Source: Online Published, Tripura, India]
Figure 4: Usage of the proposed character Chakma Dependent Vowel Sign LHAA of Chakma [from Prescribed as a Text Book by the Secretary District School Education Board, CADC, India. vide Notification No. 26/TMSTB/2008-2009/84/200 Dt. 12/1/09.]
Figure 5: Usage of the proposed character Chakma Dependent Vowel Sign AA of Chakma [from Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011). Source: Online Published from Tripura, India]
Figure 6: Usage of the proposed character Chakma Dependent Vowel Sign AA of Chakma [from Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011). Source: Online Published from Tripura, India]
Figure 7: Usage of the proposed character Chakma Dependent Vowel Sign AA of Chakma [from 30 Minide Chakma Legha Shigho (Learn Chakma Script within 30 minutes)]
Figure 8: Usage of the proposed character Chakma Dependent Vowel Sign AA of Chakma [from Prescribed as a Text Book by the Secretary District School Education Board, CADC, India. vide Notification No. 26/TMSTB/2008- 2009/84/200 Dt. 12/1/09.]
Figure 9: Usage of the proposed character Chakma Dependent Vowel Sign EI of Chakma [from Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011). Source: Online Published from Tripura, India]
Figure 10: Usage of the proposed character Chakma Dependent Vowel Sign ই of Chakma [from 30 Minide Chakma Legha Shigho (Learn Chakma Script within 30 minutes)]
Figure 11: Usage of the proposed character Chakma Dependent Vowel Sign EI of Chakma [from Prescribed as a Text Book by the Secretary District School Education Board, CADC, India. vide Notification No. 26/TMSTB/2008-2009/84/200 Dt. 12/1/09.
Figure 12: Usage of the proposed character Chakma Dependent Vowel Sign $AA$ & $EE$ of Chakma [from Chakma Legha Shigheni Bozar 2011 (Chakma Script learning year-2011). Source: Online Published from Tripura, India]
Figure 13: Usage of the proposed character Chakma Dependent Vowel Sign AA & EI of Chakma [from Chandobi Baaro Mach manuscript, circa 100 years old].
Figure 14: Usage of the proposed character Chakma Dependent Vowel Sign AA & EI of Chakma [from Prescribed as a Text Book by the Secretary District School Education Board, CADC, India. vide Notification No. 26/TMSTB/2008-2009/84/200 Dt. 12/1/09.]
Figure 15: Usage of the proposed character Chakma Dependent Vowel Sign AA & EI of Chakma [from Prescribed as a Text Book by the Secretary District School Education Board, CADC, India. vide Notification No. 26/TMSTB/2008- 2009/84/200 Dt. 12/1/09.]
ISO/IEC JTC 1/SC 2/WG 2

**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS**

FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

Please fill all the sections A, B and C below.


Please ensure you are using the latest Form from [http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html](http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html) for guidance.


**A. Administrative**

1. **Title:** Proposal to encode CHAKMA LETTER LHAA, DEPENDENT VOWEL SIGNS AA & EI for Chakma

2. Requester's name: Bivuti Chakma, Andrew Glass

3. Requester type (Member body/Liaison/Individual contribution): Individual contribution

4. Submission date: 

5. Requester's reference (if applicable): 

6. Choose one of the following:
   - This is a complete proposal: Complete
   - (or) More information will be provided later: 

**B. Technical – General**

1. Choose one of the following:
   - a. This proposal is for a new script (set of characters):
     - Proposed name of script: 
   - b. The proposal is for addition of character(s) to an existing block:
     - Name of the existing block: Chakma
     - **11100-1114F**

2. Number of characters in proposal: 

3. Proposed category (select one from below - see section 2.2 of P&P document):
   - A-Contemporary
   - B.1-Specialized (small collection)
   - B.2-Specialized (large collection)
   - C-Major extinct
   - D-Attested extinct
   - E-Minor extinct
   - F-Archaic Hieroglyphic or Ideographic
   - G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided? 
   - a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document? Yes
   - b. Are the character shapes attached in a legible form suitable for review? Yes

5. Fonts related:
   - a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? Bivuti Chakma

6. **References:**
   - a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? Yes
   - b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? Yes

7. Special encoding issues:
   - Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? Yes

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at [http://www.unicode.org](http://www.unicode.org) for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C. Technical - Justification</td>
<td></td>
</tr>
<tr>
<td>1. Has this proposal for addition of character(s) been submitted before?</td>
<td>No</td>
</tr>
<tr>
<td>If YES explain</td>
<td></td>
</tr>
<tr>
<td>2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, with whom?</td>
<td>Bivuti Chakma</td>
</tr>
<tr>
<td>If YES, available relevant documents:</td>
<td></td>
</tr>
<tr>
<td>3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?</td>
<td>No</td>
</tr>
<tr>
<td>Reference:</td>
<td></td>
</tr>
<tr>
<td>4. The context of use for the proposed characters (type of use; common or rare)</td>
<td>Common</td>
</tr>
<tr>
<td>Reference:</td>
<td></td>
</tr>
<tr>
<td>5. Are the proposed characters in current use by the user community?</td>
<td>Yes</td>
</tr>
<tr>
<td>If YES, where? Reference:</td>
<td></td>
</tr>
<tr>
<td>6. After giving due considerations to the principles in the P&amp;P document must the proposed characters be entirely in the BMP?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, a rationale provided?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?</td>
<td>Yes</td>
</tr>
<tr>
<td>8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, a rationale for its inclusion provided?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, a rationale for its inclusion provided?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, a rationale for its inclusion provided?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>11. Does the proposal include use of combining characters and/or use of composite sequences?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, a rationale for such use provided?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
<tr>
<td>12. Does the proposal contain characters with any special properties such as control function or similar semantics?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, describe in detail (include attachment if necessary)</td>
<td></td>
</tr>
<tr>
<td>13. Does the proposal contain any Ideographic compatibility characters?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, are the equivalent corresponding unified ideographic characters identified?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td></td>
</tr>
</tbody>
</table>