

Proposal to Encode an Abbreviation Sign for Bengali

Srinidhi A and Sridatta A

Tumakuru, Karnataka, India

srinidhi.pinkpetals24@gmail.com, sridatta.jamadagni@gmail.com

09 July 2015

1 Introduction

This is a proposal to encode the character ° BENGALI ABBREVIATION SIGN in the Bengali block of the Universal Character Set. The character is used to write abbreviations of words in Bengali, Assamese and other languages which employ Bengali script. It is corresponding to abbreviation signs of other north Indic scripts in its structure, glyph and function, ex. Devanagari.

2 Background

The document L2/10-029 had requested to encode the Doctor sign used in Assamese language. However it was felt that it could be represented appropriately using an abbreviation sign. Anshuman Pandey in his document L2/10-050 recommended an independent encoding of an abbreviation sign for Bengali script. Abbreviation sign is independently encoded in Devanagari, Gujarati, Kaithi, Tirhuta, Newa (accepted for encoding), Khojki, Modi scripts etc. Hence, this document requests to encode an abbreviation sign similarly for Bengali script.

3 Description

The character is widely used in printed books and publications of Bengali and Assamese languages. 09BB code point is tentatively chosen as to fill up the hole in the Bengali block and not to disturb future additions of Samaveda which require continuous code points. The properties of the character in the format of the Unicode Character Database are:

```
09BB;BENGALI ABBREVIATION SIGN;Po;0;L;;;;;N;;;;;
```

4 References

Satyavrata Sāmaśrami Bhaṭṭācārya (edited), Samaveda Samhita (Kauthuma Shakha).

[সামবেদসংহিতা কৌথুমী শাখা] , Kolkata, 1885 (Saka era-1807)

महानांनार्क्षिकः ।

२७६

(অথ মহানাম্নীভাষ্যাবতরণিকা)

ঐন্দ্রীয়ং মহানাম্নীঃ শক্রীর্কৃৎ বিকর্ষিতাঃ । পঞ্চভিঃ
সহিতা অস্তে পুরীষপদনাম্ভিঃ ॥ ১ ॥ এতাঃ প্রকৃতিতন্ত্ৰিণ
উপসর্গৈস্তু সংযুতাঃ ॥ নব সঙ্খ্যা ইতি প্রাহুর্বেদাধ্যায়ন-
শালিনঃ । ঐতরেয়-ব্রাহ্মণেহপি শস্ত্রে ষোড়শি-নামকে ।
তিত্বঃ প্রোক্তা মহানাম্নাংস্ত্রৈলোক্যাংস্বত্ববর্ণনাৎ ॥ তত্র হি মহা-
নাম্নীনা মুপসর্গানুপস্থঞ্জতি ।— “অয়ং বৈ লোকঃ প্রথমঃ
মহানাম্ন্যন্তরিক্ষলোকো দ্বিতীয়াহমৌ লোকস্তু তীয়া” — ইতি ॥

•নম্বেতা: শকরীচ্ছন্দস্বা ব্যাক্রিয়ন্ত ইত্যুক্তাংশেচৎ ষট্‌পঞ্চাশ-
দক্ষরাঃ স্ত্যঃ । •তথা চ সমান্নাতঃ চতুর্বিংশত্যাদি চতুষ্কন্তর
মিতি (পি० সূ०) ॥ অস্যার্থঃ—গায়ত্রীাদীনা মতিধৃত্যন্তানাং
ছন্দসাং চতুর্বিংশত্যক্ষরাণ্যারভ্য উত্তরোত্তরং চতুৰ্ষু চতুৰ্ষু
অক্ষরেষধিকেষু সংস্থ উষ্ণিগাদিচ্ছন্দাংসি জায়ন্তে । এবং
ক্রমশোহক্ষরাধিক্যে সতি শকরী ষট্‌পঞ্চাশদক্ষরা সম্ভবতীতি ।
এতা ঋচস্ত শকরীভ্যোহধিকাক্ষরা দৃশ্যন্তে । তস্মাদাসা
মুপসর্গাক্ষরৈরাধিক্যং ন তু স্বত ইতি জ্ঞায়তে ॥

তর্হি কেপুনঃ শাকরাঃ শাদাঃ? ইত্যুচ্যতে :-

প্রথমা । “বিদামঘবহিদিয়াঃ”—ইতি দ্বিপদাশ্মপদসম্বন্ধঃ ।
 ততঃ “শিক্ষাশচীনাশ্পাতে”—ইত্যাদ্যাস্ত্রয়োহক্ষাক্ষরাঃ শাক্ষরাঃ ।
 ততঃ “স্বর্ণাংশুঃ”—ইতি পঞ্চাক্ষরঃ পাদঃ (অত্রাক্ষরবিশ্লে-
 ষেণ দ্ব্যক্ষরত্বং দৃষ্টব্যম্), “প্রচেতনপ্রচেতয়”—ইত্য-
 ক্ষাক্ষরঃ, এতৌ দ্বানুপসর্গসঞ্জে । “ইন্দ্রদ্যামায়নইষে”—

Figure 1. Excerpt from the Samaveda Samhita showing use of ABBREVIATION SIGN (from *Bhattācārya* 1885 : pp 239)

২৩. ১ ২৩ ১ ২ ৩ ১ ২ ৩ ১ ২ ৩ ১ ২ ৩
 ঐ । ইন্দ্রজ্যেষ্ঠমাতরওজিষ্ঠম্পুপুর্নিশ্রবঃ ।

১২ ২২ ৩ ১ ২ ৩ ১ ৩ ২২

যদিধ্বক্ষেমবজ্রহস্তরোদসীউভেস্থশিপ্রপপ্রাঃ ॥ ১

(ঋ. স০ ৪, ৭, ২৭, ৫)

সাঁ. ভা. প্রথম খণ্ডে—সৈযা প্রথমা । হে ‘ইন্দ্র !’
 ‘জ্যেষ্ঠ’ প্রশস্ততমম্ ‘ওজিষ্ঠম্’ অতিশ্রুয়েন বলকরম্ ‘পুপুর্নি’
 পূরকং ‘শ্রবঃ’ অম্নং ‘নঃ’ অম্বভ্যম্ ‘আতর’ আহর প্রয়চ্ছ ।
 হে ‘বজ্রহস্ত’ বজ্রবাহো ! হে ‘স্থশিপ্র’ শোভন হনুক !
 এবস্তুত ! হে ইন্দ্র ! ‘যৎ’ অম্নং ‘দিধ্বক্ষেম’ ধারয়িতুমিচ্ছেম,
 যচ্চান্নম্ ইমে পরিদৃশ্যামানে ‘উভে রোদসী’ দ্যাবাপৃথিব্যো
 ‘আ পপ্রাঃ’ আ পূরয়সি, তদনম্যাহরেত্যম্বয়ঃ । “যেনেমেচিত্র-
 বজ্রহস্ত”—ইতি বহুচাং পাঠঃ ॥ ১

সাঁ. ভা. অ. হে বজ্রহস্ত এবং সুন্দরহনুমৎ ইন্দ্র ! আমরা যে অম্ন
 ধারণ করিতে ইচ্ছা করি—এই পরিদৃশ্যামান, উর্দ্ধাধোভাগে পরিবর্তমান
 উভয়-লোক যে অম্নে পরিপূর্ণ ক্রমিতেছে, সেই অতি প্রশস্ত, বলকর ও
 আশাপূরক ঋত্বি-আম্মাদিগকে প্রদান কর ॥ ১

বা. টী. হে ‘ইন্দ্র !’ * ‘নঃ’ আমাদিগকে সেই
 জ্যেষ্ঠ, ‘ওজিষ্ঠ’, ‘শ্রবঃ’ শ্রবণীয় ‘পুপুর্নি’ আশাপূরণ—
 তৃণানিবারণ ‘আতর’ প্রদান কর ; হে ‘বজ্রহস্ত’ দৃঢ়কর্ম !
 ‘যৎ’ বাহা ‘দিধ্বক্ষেম’ আমাদের পোষণ করিতে ইচ্ছা করা
 উচিত এবং হে ‘স্থশিপ্র’ সুদর্শন ! বাহা ‘উভে রোদসী’
 এই ভূলোক হইতে দ্বালোক পর্য্যন্ত পরিপূর্ণ রাহিয়াছে ।

Figure 2. Usage of ABBREVIATION SIGN ঋ. স০ for ঋগ্বেদ সংহিতা, সাঁ. ভা. for সামবেদ ভাষ্য, বা. টী. for বাংলা টীকা (from Bhaṭṭācārya 1885: pp 162)

সাং ভাং। অথ যজ্ঞী। পূর্বন্যা যুচি প্রকৃতা হে
 ত্রাদয়ো দেবাঃ ! যুজ্জম্ 'একম্' অদ্বিতীয়ং দানকশ্মণি
 'ইমং' সোমং 'বৃষণং' কামানামভির্বর্ষকং 'কুণুত' কুরুত।
 তথা 'ইমাং' ক্রিয়াং ফলাভির্বর্ষিকাং কুরুত ॥ ৬

সাং ভাং অং। হে মিত্র প্রকৃতি (পূর্ব মন্ত্রোক্ত) দেবগণ! তোমরা
 দানকার্যে অদ্বিতীয়! এই সোমকে এবং এই ক্রিয়াকে অভিলাষবশী
 কর ॥ ৬

বাং টাং। ঈশ্বর, প্রতি ক্ষণে প্রতি স্থানে প্রতি কার্যে
 আপনাকে দেখাইয়া এই মন্ত্রটি যেন আদেশ করিতেছেন;—
 “হে প্রজাগণ! তোমরা 'একমিৎ' একমাত্র 'ইমং মাং'
 এই অমাকে 'বৃষণং' কামবর্ষা বলিয়া 'কুণুত' দৃঢ় বিশ্বাস
 কর।”

অর্থাৎ তোমাদের মঙ্গল করিতে— 'অভিলাষ পূর্ণ করিতে
 আম্ ব্যতীত আর কেহই নাই। ৬

(৭)

৩ ১ ২ ৩ ১ ২ ৩ ১ ২ ৩ ১ ২
 সনইন্দ্রায় যজ্যে বেরুণায় মরুদ্যঃ।

৩ ১ ২ ৩ ১ ২ ৩ ১ ২
 বরিবোরিৎ পরিঅব ॥ ৭

সাং ভাং। অথ সপ্তমী। গায়ত্র্যো পারমান্যো তু
 সনইত্যাদিকে ঋচৌ। অমহীযুস্তয়োরেবং ছন্দোদৈবত-
 নির্ণয়ঃ ॥ হে সোম! 'সঃ' 'বরিবোরিৎ' ধনস্য স্তকঃ স্বং
 'নঃ' অশ্বাকং যজ্যে, যজ্যায় 'ইন্দ্রায়' 'বরুণায়' 'মরুদ্যঃ'
 চ 'পরিঅব' ধারয়া ক্ষর ॥ ৭

Figure 3. Excerpt from the Samaveda Samhita showing use of ABBREVIATION SIGN (from Bhaṭṭācārya 1885: pp 170)

৯। ভা° অ° উ° আ° ৩, ১, ১০ দেখ।
* ইন্দ্রো বিরাড়িঃ; চতুর্বেহিহনি পৃষ্ঠম্।

৯। ভা° অ° উ° আ° ১, ১, ১২ দেখ।
১০। ভা° অ° উ° আ° ১, ১, ৯ দেখ।

Figure 4. Other usages of the ABBREVIATION SIGN from the Samaveda Samhita (from *Bhaṭṭācārya* 1885)

6 Proposal Summary Form

(Based on N4502-F)

A. Administrative

1. Title:	Proposal to Encode an Abbreviation Sign for Bengali	
2. Requester's name:	Srinidhi A and Sridatta A	
3. Requester type (Member body/Liaison/Individual contribution):	Individual Contribution	
4. Submission date:	2015 July 09	
5. Requester's reference (if applicable):		
6. Choose one of the following:		
This is a complete proposal:	Yes	
(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):	No				
Proposed name of script:					
b. The proposal is for addition of character(s) to an existing block:	Yes				
Name of the existing block:	Bengali				
2. Number of characters in proposal:	1				
3. Proposed category (select one from below - see section 2.2 of P&P document):					
A-Contemporary	<input checked="" type="checkbox"/>	B.1-Specialized (small collection)	<input type="checkbox"/>	B.2-Specialized (large collection)	<input type="checkbox"/>
C-Major extinct	<input type="checkbox"/>	D-Attested extinct	<input type="checkbox"/>	E-Minor extinct	<input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic	<input type="checkbox"/>	G-Obscure or questionable usage symbols	<input type="checkbox"/>		
4. Is a repertoire including character names provided?	Yes				
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?					
	Not necessary. The required glyph may be copied from Devanagari.				
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):					

Not necessary.

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; see proposal for additional details.

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> 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.

See the detailed proposal

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	No
If YES explain	
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.)?	No
If YES, with whom?	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	
Reference:	Yes
4. The context of use for the proposed characters (type of use; common or rare)	
Reference:	The character is in common use.
5. Are the proposed characters in current use by the user community?	Yes
If YES, where? Reference:	The character is used by the Bengali and Assamese communities
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	Yes
If YES, is a rationale provided?	Yes
If YES, reference:	The Bengali block is in the BMP
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	N/A
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	No
If YES, is a rationale for its inclusion provided?	N/A
If YES, reference:	N/A
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	N/A
If YES, reference:	N/A
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?	Yes
If YES, is a rationale for its inclusion provided?	Yes

If YES, reference:	See text of proposal	
11. Does the proposal include use of combining characters and/or use of composite sequences?	No	
If YES, is a rationale for such use provided?	N/A	
If YES, reference:		
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	N/A	
If YES, reference:	N/A	
12. Does the proposal contain characters with any special properties such as		
control function or similar semantics?	No	
If YES, describe in detail (include attachment if necessary)	N/A	
13. Does the proposal contain any Ideographic compatibility characters?	No	
If YES, are the equivalent corresponding unified ideographic characters identified?	N/A	
If YES, reference:	N/A	