L2/05-384

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps.***

A. Administrative

1. **Title:** *Proposed Changes to Gurmukhi 3*
2. Requester's name: *Sukhjinder Sidhu (Punjabi Computing Resource Centre), Unicode Technical Committee*
3. Requester type (Member body/Liaison/Individual contribution): *Individual*
4. Submission date: *2005-12-28*
5. Requester's reference (if applicable): *L2/05-344*
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):
        Proposed name of script:
    b. The proposal is for addition of character(s) to an existing block: *Yes*
        Name of the existing block: *Gurmukhi*
2. Number of characters in proposal: *1*
3. Proposed category (select one from below - see section 2.2 of P&P document):
    A-Contemporary   **X**   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. Proposed Level of Implementation (1, 2 or 3) (see Annex K in P&P document): *2*
    Is a rationale provided for the choice? *Yes*
        If Yes, reference: *Combining Mark*
5. 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*
6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for
    publishing the standard? *Dr Kulbir Thind (AnmolUniBani, OpenType).*
    If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools
    used: *E-mail sukhuk@users.sourceforge.net*
7. 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*
8. 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*
   *Sorting*

9. 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 http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

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.)?	*Yes*
If YES, with whom?	*Jeevan Deol, Manudeep Singh, Serjinder Singh, Kulbir Thind and general community consultation*
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?	*No*
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	*Rare*
Reference:	
5. Are the proposed characters in current use by the user community?	*Yes*
If YES, where? Reference:	*Used in the Adi Granth, the holy book of the Sikhs*
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:	*In keeping with the Gurmukhi block*
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	
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?	
If YES, reference:	
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?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?	*Yes*
If YES, is a rationale for its inclusion provided?	*Yes*
If YES, reference:	*See End*
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?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
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)	
13. Does the proposal contain any Ideographic compatibility character(s)?	*No*
If YES, is the equivalent corresponding unified ideographic character(s) identified?	
If YES, reference:	
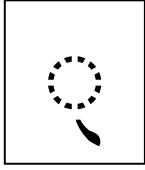
U+0A51   GURMUKHI SIGN UDAAT

*indicates a high tone

GURMUKHI SIGN UDAAT is used in the Adi Granth, the holy book of the Sikhs, to indicate a high tone. It is no longer used in modern Punjabi, and instead, differing tones are indicated by the use of aspirate consonants and the subjoined form of /ha/.

The sign is similar in appearance to GURMUKHI SIGN VIRAMA. The Virama sign was invented for Gurmukhi so that it could fit in with the Virama based model used by all the other Indic scripts. Virama is **not** a sign traditionally used in Gurmukhi. The behaviour of Udaat is radically different and requires separate representation.

It is recommended that Udaat be encoded as a separate Unicode character, with the following properties:

```
0A51;GURMUKHI SIGN UDAAT;Mn;0;NSM;;;;;N;;;;;
```

For the purposes of collation, GURMUKHI SIGN UDAAT should sort as if it is the subjoined form of /ha/. The existing Unicode Collation Charts should be altered to:

	ਸ	ਹ	ੑ	ਕ	ਖ	
…	0A38	0A39	0A51	0A15	0A16	…

Udaat should be placed after the consonant whose tone is being changed but before the vowel. In many ways, **Udaat should be treated as a subjoined consonant**. In the following examples, an acute accent indicates the high tone.

ਖੋਲ੍ਇਓ (Khōlí 'ō)

0A16	0A4B	0A32	0A51	0A3F	0A13
ਖ	ੋ	ਲ	ੑ	ਿ	ਓ

ਸੰਮ੍ਾਲੇਹਾਂ (Samhálēhāṁ)

0A38	0A70	0A2E	0A51	0A3E	0A32	0A47	0A39	0A3E	0A02
ਸ	ੰ	ਮ	ੑ	ਾ	ਲ	ੇ	ਹ	ਾ	ਂ

ਓਲਾਮ੍ੇ (Ōlāmhé)

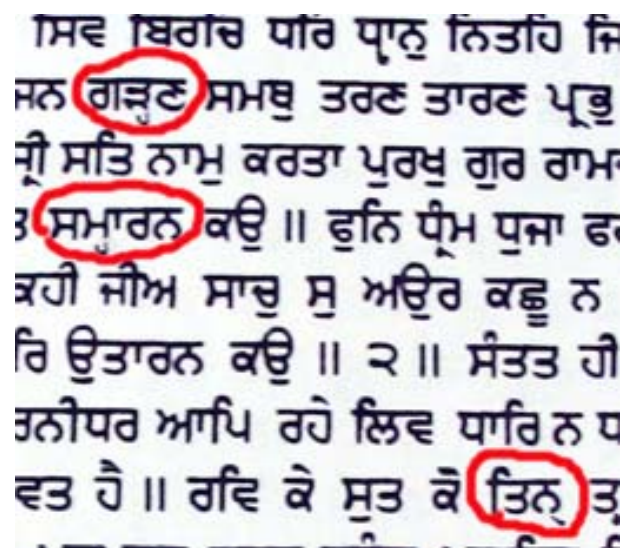| 0A13	0A32	0A3E	0A2E	0A51	0A47
ਓ	ਲ	ਾ	ਮ	ੑ	ੇ

When used with GURMUKHI SIGN U and GURMUKHI SIGN UU, the vowels should be pushed below Udaat. This is how subjoined consonants are treated now in Gurmukhi.

Usage

Evidence for its use is provided:
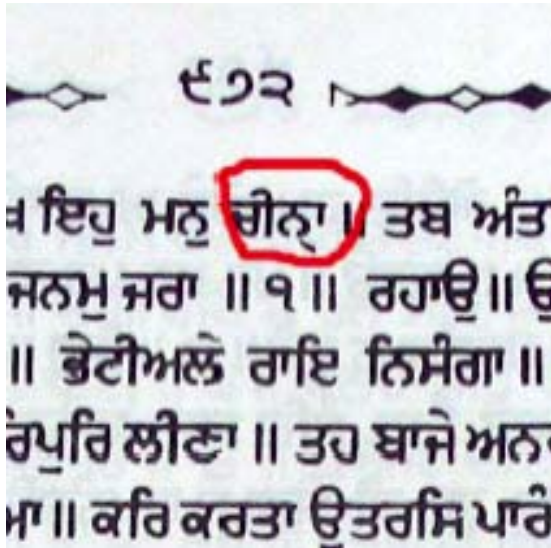
1 Guru Granth Sahib page 972 (੯੭੨)   2 Guru Granth Sahib page 1404 (੧੪੦੪)

Background Information Submitted to the UTC

Udaat[2] looks like the Halant or Virama character in Devanagari, but it is not that character. It is found in the Sri Guru Granth Sahib 1188 times[3]. The Udaat is/was used for a non-segmental phoneme (*akhāndi tòni*) known as the high tone[4]. This sign is related to Ha, because Ha itself is used to distinguish tones, but it is not a variant form. Udaat is related to Devanagari Udatta (U+0951) which also indicates a high tone in Sanskrit literature (however its technical properties differ).

High tone is still present in modern Punjabi, however, the Udaat is not used in modern Gurmukhi. In modern Gurmukhi, there are no symbols that highlight the high or low tones. But at places where the Udaat was used earlier, now another symbol known as the Pairin Haha is being used. This does not mean that the Udaat is equivalent to Pairin Haha. The orthographical rules of Gurmukhi suggest that Pairin Haha is used for the pronunciation of an aspirated sound of the initial letter[5]. However, in various Punjabi dialects, we find a variety of pronunciations, such as in the Majhi of Central Punjab, the words written with Pairin Haha would certainly be pronounced with a high tone, however, in most other dialects (both Western Punjabi and Eastern Punjabi dialects), either a complete or seminal /h/ would be found, or in places we would find the aspirate sound[6].

In the Old Gurmukhi of the Sri Guru Granth Sahib, both Udaat and Pairin Haha have been used. This is a result of the wide range of Punjabi dialects, apart from other languages, being represented in Gurbani, at different stages in their evolution (from the 12th century to the 17th century). The Udaat suggests the high tone, while the Pairin Haha denotes the aspirate /h/ with the inherent vowel being suppressed. In modern Gurmukhi, only Pairin Haha is used, but orthographically it does not suggest the high tone. Both high tone and /h/ pronunciation are to be found among the Punjabi dialects.

---

[2] The *Punjabi-English dictionary*, published by Punjabi University, Patiala (1994) gives following meanings of the term Udaat: 'sublime; acutely accentuated, sharply intoned' (p. 9)
[3] Kulbir S Thind, *Text Trivia* in Gurbani-CD 2004. The basis of the file is the Sri Guru Granth Sahib, published by the Shiromani Gurdwara Parbandak Committee in 1994.
[4] Harkirat Singh, *Gurbani di Bhasha te Vyakaran* (1997, in Punjabi), pp. 102-3.
[5] Joginder Singh Talwara, *Gurbani da Saral Viakarn-Bodh*, part I, pp. 27-8.
[6] Ibid, p. 103.