ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

A. Administrative

1	Title:	Proposal to add Six characters in the Devanagari block for representation of Kashmiri language in Devanagari script.
2	Requester's name:	Ministry of Communications & IT, Government of India Through: Bureau of Indian Standards
3	Requester type (Member body/Liaison/Individual contribution):	National Member Body
4	Submission date:	14-10-2009
5	Requester's reference (if applicable):	
6	Choose one of the following: This is a complete proposal: or, More information will be provided later:	Yes If required more information may 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:	No
	b. The proposal is for addition of character(s) to an existing block: Name of the existing block:	Yes DEVANAGARI
2	Number of characters in proposal:	Six
3	Proposed category: 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	A
4	Proposed Level of Implementation (1, 2 or 3): Is a rationale provided for the choice? If Yes, reference:	1
5	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? 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? If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:	Government of India

7	References: Are references (to other character sets, dictionaries, descriptive texts etc.) provided? 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)?	No
9	Additional Information:	

C. Technical - Justification		
1	Has this proposal for addition of character(s) been submitted before? If YES explain	No
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.)? If YES, with whom? If YES, available relevant documents:	Yes User Groups
3	Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	20 Lakhs approximately
4	The context of use for the proposed characters (type of use; common or rare): Reference:	Common
5	Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes Kashmir region of India as well as Kashmiri spread across India & world
6	After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	Yes
7	Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Preferably these should be placed contiguous
8	Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9	Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No

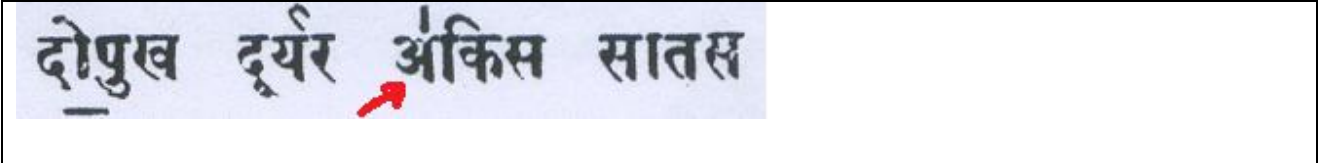| 10	Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	No
11	Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a 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:	No
12	Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	No
13	Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:	No

Kashmiri in Devanagari script

Kashmiri is mainly written in Perso-Arabic. Devanagari script is also used for writing Kashmiri. Central Hindi Directorate (CHD), Ministry of Human Resources Development, Government of India is the recommending authority of *Parivardhit Devanagari* (Extended Devanagari Set for representation of the other Indian languages in Devanagari). Department of Information Technology and CHD in consultation with all the stakeholders of the Kashmiri language which includes Experts from CIIL Mysore, Central Hindi Directorate (CHD), Commission for Scientific & Technical Terminology (Ministry of MHRD, Govt. of India), C-DAC and other Kashmiri-Devanagari experts finalized the additional characters which need to be encoded in the ISO 10646 and Unicode. The experts opined that there is need to encode 10 characters in the Unicode/ ISO 10646 for representation of the Kashmiri language. These characters are shown in the table below:

	Glyph	Name
1	अ़	Kashmiri Devanagari Vowel High Central Unrounded U
2	अ़	Kashmiri Devanagari Vowel High Central Unrounded UU
3	◌	Kashmiri Devanagari Vowel Sign High Central Unrounded U
4	◌	Kashmiri Devanagari Vowel Sign High Central Unrounded UU
5	अ	Kashmiri Devanagari Vowel Mid Central Unrounded O
6	आ	Kashmiri Devanagari Vowel Mid Central Unrounded OO
7	◌॔	Kashmiri Devanagari Vowel Sign Mid Central Unrounded O
8	◌ा	Kashmiri Devanagari Vowel Sign Mid Central Unrounded OO
9	औ	Kashmiri Devanagari Vowel Back Low Rounded O
10	◌ौ	Kashmiri Devanagari Vowel Sign Back Low Rounded O

The characters shown at Serial No. 1-4 are already proposed through document no. ISO/IEC JTC1/SC2/WG2 N3480 and L2/08-250 by Mr. Praveen Satpute and Mr. Micheal Everson. Government of India has already recommended the encoding of these four characters through letter no. 13(6)/2008 -HCC(TDIL) dated August 11th, 2009 submitted to Unicode Consortium. The reference document no. is L2/09-311, as placed at Annexure-I.

The characters shown at Serial No. 5, अॅ is Devanagari Vowel Mid Central Unrounded O and character at Serial No. 7, ◌ॅ is the Vowel Sign for the same. This character is very frequently used in the Kashmiri Devanagari text. Following are some published examples of the usage of these characters.

Example 1 - Usage of "Devanagari Vowel Mid Central Unrounded O" taken from "Ramayan Kashmiri Bhasha Mein" shown with red arrow.

Example 2 - Usage of "Devanagari Vowel Sign Mid Central Unrounded O" taken from "Ramayan Kashmiri Bhasha Mein" shown with red arrow.

The characters shown at Serial No. 6, आॅ is "Devanagari Vowel Mid Central Unrounded OO" and character at Serial No. 8, ◌ाॅ is the Vowel Sign for the same. This character is very frequently used in the Kashmiri Devanagari text. Following are some published examples of the usage of these characters.

Example 3 - Usage of "Devanagari Vowel Mid Central Unrounded OO" taken from "Ramayan Kashmiri Bhasha Mein" shown with red arrow.

Example 4 - Usage of "Devanagari Vowel Sign Mid Central Unrounded OO" taken from "Ramayan -Kashmiri Bhasha Mein" shown with red arrow.

The complete text of the reference page may be seen at Annexure -II. There are many examples of usage of these characters in different publications as shown below:

ॐ अक्षरसंकेतपरिपाटीयम् ।

तत्रादौ काश्मीरभाषायां वर्गचतुर्थाक्षराणि क्वचिच्चोच्चार्यन्ते यकारश्च
प्रसिद्धोच्चारणतया न शब्दायत इति बोध्यम् ॥

प्रसिद्धा वर्णाः ।

अ आ इ ई उ ऊ ऋ ॠ ऌ ॡ ए ऐ ओ औ अं अः क ख ग ङ च छ
ज ट ठ ड त थ द न प फ ब म य र ल व श ष स ह क्ष ॥

अप्रसिद्धानां संकेतः ।

अं आं इं ऊं ऋं च छ ज झ ॥

भाषाशब्दाः ।	अगुनु ।	अलतु ।	आशिद् ।	एंजुलुं ।	ऋथ् ॥
संस्कृताः शब्दाः ।	अग्निः ।	अलक्तः ।	आश्विनः ।	उज्ज्वलः ।	ऋतुः ॥
भाषाशब्दाः ।	पंपु ।	तंतु ।	दांरू ।	गुहैं [?] ।	हिरू ।	कैंमू ॥
संस्कृताः शब्दाः ।	पक्वः ।	तनु ।	इमस्तु ।	अभ्रः ।	शीतम् ।	कृमिः ॥
भाषाशब्दाः ।	वेंरि ।	पंकि ।	देंरि ॥	कंहू ।	पंपु ।	देंह ॥
संस्कृताः शब्दाः ।	वलयाः ।	पक्वाः ।	दृढाः ॥	वलयः ।	पक्वः ।	दृढः ॥
भाषाशब्दाः ।	देंरू । पंपू । कंरू ॥	चंहू । छेंहू । जंहू । पोङ् ॥
संस्कृताः शब्दाः ।	दृढा । पक्वा । अमसूता ॥	बहु । रिक्तम् । बधिरः । पानीयम् ॥

अत्रोर्ध्वचिह्नं स्वरस्याधोविन्दुर्व्यञ्जनस्य ज्ञेया । तथा ऊकारमात्रा (पू)

Usage Example 5: The Kashmiri Vowels shown in "KAshmIra-shabdAmRtaM" written by Shri Ishwar Kaul and edited by Mr. George Grierson, published in 1889.

THE KĀSHMĪRĪ ALPHABET IN THE ROMAN CHARACTER WITH

A. VOWELS

Letter	Nāgarī Equivalent	Persian Equivalent	As in
a	˘ (medial only)	´	tᵃh, चुह्, چِه
å	˘ (medial only)	´	andᵃrⁱ, चन्द्रिर्, اَنْدَرِ
a	अ	۱, ´	abal, अबल्, أَبَل
ạ	ब, ॒	۱, ´	aḍᵃrun, बद्रन्, أَدَرُن ; ganzᵃr, गंज़र, گَنَزَر
à	ब, ॒	۱, ´	àsⁱ, बसि, أَس ; kàrⁱ, करि, کَرِ
ā	आ, ा	آ	āb, आब्, آب ; kāṭh, काठ्, کاتھ
ai	ऐ, ै	ـِي ,أَي	aiṭh, ऐठ्, أَيـِٹھ ; mail, मैल्, مَيل
au	औ, ौ	ـُو	gauv, गौव्, گَوو
ă	˘	̆	bañᵉrāwun, बंज्रावुन्, بِنزراوُن
ĕ	य (after a consonant)	̆	vĕth, व्यथ्, وِتھ ; bĕñĕ, ब्यञ, بِنِه. Initial य, or य after a vowel, is yĕ, as in yĕmis, यमिस्, يَمِس ; biyĕ, बिय, بِيِه
	or ब (after certain consonants)		
ẹ̆	य (after a consonant)	̅	vẹ̆thᵃrun, ब्दरन्, وِتھَرُن
	or बं (after certain consonants)		
ĕ	ditto	̅	ṭĕ̆thᵘ, ट्दू, ٹِتھ
ē	˗ (never initial)	ي	mēl, मेल्, ميل
i	ि ,, ,,	̅	àsⁱ, बसि, أَس
i	ि ,, ,,	̅	nishĕ, निश, نِشِه
ī	ी ,, ,,	ـِي	shīn, शीन्, شِين
o	॒ (medial only)	´	dᵒnᵘ, दुनु, دُن
ŏ	व (after a consonant)	̈ .	dŏd, द्द्, دُد
ọ̆	व ,, ,,	´	gọ̆nᵘ, ग्वनु, گُن
ŏ	व ,, ,,	´	chŏpᵘ, छुपू, چِھُپ
o	ब, ॒	´	ogun, ब्गुन्, أُگُن ; bodᵘ, बड्, بُد
ō	ओ, ो	و ,او	ōr, ओर्, اور ; pōsh, पोश्, پوش
ô	ओ, ो	و ,او	ôsᵘ, ओमु, اوس ; môlᵘ, मोलु, مول
ö	आ, ॒	آ (initial and medial)	ös, आस, آس ; mölⁱ, मालि, مآل
u	॒ (never initial)	omitted	korᵘ, करु, کُر
u	॒ ,, ,,	´	kus, कुस्, کُس
ū	॒ ,, ,,	ـُو	ṭūr, चूर, چُور
ü	॒ or ॒ ,, ,,	omitted, or ̅	kürᵘ, कर, کَر ; abadᵘrᵘ, अबद्र, أَبَدَر
ü	ब, ॒	أ, ´	üñᵘ, ब्ञु, أَن ; büḍᵘ, बड्, بُد
ü	क, ॒	ي	süty, सूथ्, سيت

The nasalization indicated in the Nāgarī character by *anunāsika* (ँ) is represented in the roman character by the sign ~ over the nasalized vowel. Thus आँ ā̃, and so on. This does not affect the alphabetical order. *Anusvāra* (ं) is represented by m or n according to its sound. In Kāshmīrī words it occurs only as a *compendium scripturæ* for a nasal before a consonant of the same class. Thus, स्वंड् for स्वण्ड् mọ̆ndᵘ.

Usage Example 6: The Kashmiri Vowels shown in Grierson Dictionary.

Other common words examples for the proposed characters:

अं
◌ं

Arm -> नंर

bird -> च्रंर

unbroken -> अंर

butter -> थंन्य

आं
◌ां

mother -> मांज

(she) was -> आंस्य

last -> आंखरी

to unknown -> गांरन

boatman -> हांन्ज़

The characters shown at Serial No. 9, औ is "Devanagari Vowel Back Low Rounded O" and character at Serial No. 10, ौ is the Vowel Sign for the same. This character and its vowel sign are recommended by the experts to represent Kashmiri sound close to pronunciation of International Phonetic symbol ɔ. Presently it is represented by the consonantal alphabet व (va). All the Devanagari vowels have vowel signs (*matras*) but the Devanagari consonantal alphabet letter व (va) does not have any vowel sign. Hence the experts have recommended new signs (glyphs) for this sound and its vowel symbol.

Word Examples:

Hindi word सुख (Sukh) is pronunced as sɔkh in Kashmiri, if the present व (va) is used for writing this word it will be written as स्वख and a person not knowing Kashmiri will pronunce it as "svakh", which is not correct. But when it is written using the new vowel sign it will be pronounced correctly i.e. सुख -> sukh -> sɔkh -> सौख.

Other example words:

खुश -> khush -> khɔsh -> खौश (happy)

पैर -> khush -> khɔsh -> खौर (foot)

दूध -> doodh -> dɔd -> दौद (milk)

Kashmiri words like drɔkh (vomit), nyɔth (thumb) etc can not be written at all with the help of व (va) as it is not having any vowel sign (*matra*). But these may be correctly represented using new signs.

vomit -> द्रौख -> drɔkh

thumb -> न्यौठ -> nyɔth

The printed usage examples for these characters are not presently available as the experts have recommended its usage in all future publications of Kashmiri language text in Devanagari script. CHD has also recommended usage of these characters and these will be shortly recommended in the official publication of Central Hindi Directorate *"Devanagari Lipi Tatha Hindi Vartani Ka Mankikaran"*.

L2/09-311

Swaran Lata
Director & HoD
Human Centred Computer Division

भारत सरकार
GOVERNMENT OF INDIA
संचार और सूचना प्रौद्योगिकी मंत्रालय
MINISTRY OF COMMUNICATIONS AND INFORMATION TECHNOLOGY
सूचना प्रौद्योगिकी विभाग
DEPARTMENT OF INFORMATION TECHNOLOGY
इलेक्ट्रॉनिक्स निकेतन
ELECTRONICS NIKETAN
6, सी.जी.ओ. कॉम्प्लेक्स/6, C.G.O. COMPLEX
नई दिल्ली/New Delhi-110003
दिनांक/Dated............

Email: slata@mit.gov.in
Telephone: 011-24301242

दूरभाष/Tele :
संलग्न पत्र सं० :
D.O.No...... 13(6)/2008-HCC(TDIL)

August 11th, 2009

Subject: Encoding of additional characters in Unicode for representation of Kashmiri.

Dear Dr Davis,

This is with reference to the encoding of the additional characters for representation of Kashmiri in the Unicode Standard. The issue was discussed by the experts in detail, during various meetings at DIT and Central Hindi Directorate (CHD), New Delhi.

Department of IT, Govt of India has finalized the additional characters required for Kashmiri in consultation with CHD. There is need to encode 10 characters in the Unicode for the same. Out of these 10 characters, 4 characters are already proposed through Document L2/08-250 (ISO/IEC JTC1/SC2/WG2 N3480). Department of IT recommends the encoding of these four characters in the Devanagari block of Unicode Standard with the GOI recommended names as shown in the attached document. For the remaining six characters DIT will be submitting the formal proposal by September end.
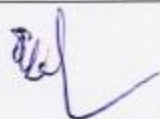
Thanking You.

Yours sincerely,

(Swaran Lata)

Dr. Mark E. Davis,
President
The Unicode Consortium
P.O. Box 391476
Mountain View, CA 94039-1476
USA

Kashmiri Vowels

	Code	Glyph	Name
1		अ॒	Devanagari Vowel High Central Unrounded U
2		अ॒	Devanagari Vowel High Central Unrounded UU
3		○॒	Devanagari Vowel Sign High Central Unrounded U
4		○॒	Devanagari Vowel Sign High Central Unrounded UU

रामायण

कश्मीरी भाषा में

Edited and Published

BY

P. DAMOODHAR KOUL

REVISED BY

P. SHYAMLAL BIRA

SAMBAT BIKRAMI 2000

दामोदर कोल मैनेजर के प्रबन्ध से
काश्मीर सटेंडार्ड प्रेस श्रीनगर में छपी

मु वद बुञ वदन स्त्य गोय रंग बेरंग   मु वद बुञ वदन सत्य श्रीश्स प्यवुय संग
वोलुख थलि बुरज्ञ त्रोबुख खासु मलमल   पकान गथ त्रनबय अज राहि जंगल
ति याञ बुछ शेहर क्यव लोकव रिवान द्राय   दोपुख क्या सन वनस मंज कति रटन जाय
दिलस प्यठ दाग रोट बुजुलयव गुलालव   दोपुख दूर्यर अंकिस सातस न चालव
गुलाबस सोसनुक ह्रुन रंग साम्पुञ   ति डीशित हाल मसवल ह्रचन कोम्पुन
सपुञ सारी प्रजलबुञ गुल अवार   फोलन त्यलि थलि दर्शन दिन ढुबार
पकान गय तथ कोहस प्यठ आल्यहिमुत रंग   बदल गव ज्यंठ पोशस कारतिक कुंग
सु कीकी शीन छट चंज मंजहूर्य गंयि तेज   बनस कुन लेज खलांयिकन पोह पनस रेज
चटुख मंजिल रटुख मंज बन खोटुख पान   खलुख पथ फीर सारी आय रिवान
तिथय गय तिम डन्डक बन मंज रटुख जाय   जन्म क्रेछर तु कर्मस क्या छु परवाय
खबर थलि बूज कौसल्यायि कोत गव   बननि लेज राज पोत्रस कुन चु कन थव

—∞∞∞—

राजा राम चन्द्र जी के फिराक में कौसल्या का बिलाप करना

कौसल्यायि हुन्दि गोबुरो ॥ करयो गूर गूर ॥ परांयो राम रामां ॥ करयो गूर गूर कोतू गोहं चु त्रोबिथ ॥ कसू हक हाल बांविथ ॥ अनी कुस मन नांविथ ॥ करयो० लगुयो पोत छाये ॥ हिय कंरथस बु हाये ॥ नारस बुट बु लाये ॥ करयो० ॥ कनन छय कनु वाजे ॥ श्री कृष्ण महाराजे ॥ जगतुक छुख चु राजे ॥ करयो० ॥ फेरहय अन्दि अन्दी लागय पोशि गन्दी ॥ जामु चान्य सुन सुन्दी ॥ करयो० ॥ लोल चोन बुञ म्ये आमो ॥ छारथो शेहर गामो ॥ ब्ययि गूछहम चु रामो ॥ करयो० ॥ नेरहा बाजारी ॥ ह्रडनम लूख सारी ॥ पांदन लगय पारी ॥ करयो० ॥ नेरुयो दारिपती ॥ लागुयो कारिपती ॥ तारि दिल गोम यंती ॥ करयो० ॥ म्ये कमो शाफ आंसी ॥ तिम ति कोन कांसि कांसी ॥ चु गोहम बन वांसी ॥ करयो० ॥ च्यं प्रथम बुरज जामय ॥ बु छारथ शेहर गामय ॥ परयो राम राम ॥ कारयो० ॥ लुलि मंज ललु नावथ ॥ जिगरस मंज बु सावथ बुञ तिनो कांसि हावथ ॥ करयो० ॥ नेरुयो शाम लटे बार म्यांज छिय च्यं मटे ॥ गाश्र लाल त्रटे ॥ करयो० ॥ दूर्यर नो बु चालय ॥ कसू कंरथस इवाळय ॥ लांजथस मांपि

भारत सरकार
GOVERNMENT OF INDIA
केंद्रीय हिंदी निदेशालय
CENTRAL HINDI DIRECTORATE
मानव संसाधन विकास मंत्रालय
(MINISTRY OF HUMAN RESOURCE DEVELOPMENT)
(उच्चतर शिक्षा विभाग)
(DEPARTMENT OF HIGHER EDUCATION)
पश्चिमी खंड–7, रामकृष्णपुरम, WEST BLOCK-VII, RAMA KRISHNA PURAM
नई दिल्ली/NEW DELHI-110066   दिनांक/DATED :

9.7.2009

सेवा में

निदेशक
संचार और सूचना प्रौद्योगिकी मंत्रालय
सूचना प्रौद्योगिकी विभाग
इलेक्ट्रॉनिक्स निकेतन
6, सी.जी.ओ.कॉम्प्लेक्स
नई दिल्ली-110003

महोदय,

केंद्रीय हिंदी निदेशालय द्वारा आयोजित बैठक में कश्मीरी भाषा के विशेषज्ञों ने जो चिह्न निर्धारित किए हैं उसकी एक प्रति आपको सूचना हेतु भेजी जा रही है।

शुभकामनाओं सहित,

भवदीया

(डॉ.शशि भारद्वाज)
उपनिदेशक

संलग्न - कश्मीरी के लिए निर्धारित नागरी चिह्न

Please file Unicode for Kashmiri-Devanagari

TDIL (Section)

18/8/2009

कश्मीरी भाषा के लिए देवनागरी लिपि चिंतन

स्वर - अ आ इ ई उ ऊ ऋ ए

ा ि ी ‍ु ‍ू ‍ृ े

(ऐ) औ (औ) (अं) (अः)

(ऐ) ो (ै) ‍ो (ं) (ः)

कश्मीरी के विशेष स्वर - अ॒ आ॒ अ॑ औ॑ ऊ॑/ओ॑ (...) ऐ॑ ओ॑

(कश्मीरी के विशेष स्वरों का कोशीय क्रम 'ऊ' के पश्चात् होगा ।)

व्यंजन - क ख ग (घ) (ङ)

च छ ज (झ) ञ

च॒ छ॒ ज॒

ट ठ ड (ढ) (ण)

त थ द (ध) न

प फ ब (भ) म

य र ल व (ड़)(ढ़)

श ष स ज़ ह

(क्ष) (ज्ञ) (त्र) (श्र)

नोट : 1. कोष्ठक में रखे गए वर्ण आगत अथवा तत्सम शब्दों के लिए ही सुरक्षित हैं, मूल कश्मीरी शब्दों में ये वर्ण प्रयुक्त नहीं होते ।

2. कश्मीर मूल शब्दों में घोष महाप्राण ध्वनियों का अभाव है । तत्सम तथा आगम शब्दों में इनका प्रयोग संभव है ।

3. ऋँ

(श्री बी. एन. बेताब)

(प्रो. मोहनलाल स)

18/6/2009
(डॉ. शशिशेखर तोषखानी)