> Universal Multiple-Octet Coded Character Set International Organization for Standardization Organisation Internationale de Normalisation Международная организация по стандартизации

## Doc Type: Working Group Document

Title: Proposal to encode 10 Latin letters for pre-1921 Latvian orthography
Source: German NB
Status: National Body Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2009-03-26

## Additions for pre-1921 Latvian orthography

## G U+A7AO LATIN CAPITAL LETTER G WITH OBLIQUE STROKE

- U U+A7A1 LATIN SMALL LETTER G WITH OBLIQUE STROKE

K U+A7A2 LATIN CAPITAL LETTER K WITH OBLIQUE STROKE
k U+A7A3 LATIN SMALL LETTER K WITH OBLIQUE STROKE
Af U+A7A4 LATIN CAPITAL LETTER N WITH OBLIQUE STROKE
A U+A7A5 LATIN SMALL LETTER N WITH OBLIQUE STROKE
R U+A7A6 LATIN CAPITAL LETTER R WITH OBLIQUE STROKE
f U+A7A7 LATIN SMALL LETTER R WITH OBLIQUE STROKE
\& U+A7A8 LATIN CAPITAL LETTER S WITH OBLIQUE STROKE - Also used in pre-1950 Lower Sorbian orthography
\& U+A7A9 LATIN SMALL LETTER S WITH OBLIQUE STROKE
$\rightarrow$ A79E latin small letter long s with diagonal stroke
Informative notes and cross references to be added for existing characters:
U+1E9C LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE $\rightarrow$ A7A9 latin small letter s with oblique stroke

- Medievalist use for several abbreviations
- Used in pre-1921 Latvian orthography
- Used in Blackletter types of pre-1950 Lower Sorbian orthography

All these letters are used in the orthography of Latvia used before 1921 [4] [5] [7], see fig. 1 to 6,9 to 11. The LATIN SMALL LETTER S WITH OBLIQUE STROKE also occurs in a long s form; this form is unified with the already encoded U+1E9C LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE.
Also, the stroked forms of L/I which are part of the pre-1921 Latvian orthography are proposed to be unified with the already encoded U+0141/0142 LATIN CAPITAL/SMALL LETTER WITH STROKE.
The encoding of the letters proposed here closes a gap regarding Blackletter (Fraktur) types:
According to [4] (which appeared 1880) pp. 227-234, Blackletter was in common use during the 19th century for the following languages: German, Danish, Swedish, Latvian, Finnish, Czech, Slovak. Even as this list may be incomplete, all letters used for the 19th century orthographies of all major European languages are now included in Unicode, with the exception of Latvian.
Encoding the 10 letters proposed here would enable to create Blackletter fonts which cover the complete 19th century and early 20th century usage in Europe.
The stroked forms of $\mathrm{G} / \mathrm{g}, \mathrm{K} / \mathrm{k}, \mathrm{L} / \mathrm{l}, \mathrm{N} / \mathrm{n}, \mathrm{R} / \mathrm{r}$ (but not the stroked form of $\mathrm{S} / \mathrm{s} / \mathrm{I}$ ) roughly conform to the modern use of cedilla in the modern Latvian orthography (which is frequently displayed like a comma below for these members, or as a turned comma above in the case of the lowercase g). However, the stroke cannot be simply regarded as a typographic variant of the modern cedilla, as its appearance is
completely different not only in Blackletter (which was the prevalent printing type before 1921 in Latvia), but also in Roman type (see fig. 1, 9 to 11).
Besides the historic Latvian use, the s with oblique stroke also was used in Lower Sorbian orthography until about 1950 (until the current orthography was introduced) [6]. - See fig. 7 and 8.

Regarding the representative glyphs, the stroke is presented in a flat angle, according to the glyphs presented in fig. 1. Therefore, in the names the term "oblique stroke" is used regarding the stroke.

## Properties:

```
A7A0;LATIN CAPITAL LETTER G WITH OBLIQUE STROKE;Lu;O;L;;;;;N;;;;A7A1;
A7A1;LATIN SMALL LETTER G WITH OBLIQUE STROKE;Ll;0;L;;;;;N;;;A7A0;;A7A0
A7A2;LATIN CAPITAL LETTER K WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A3;
A7A3;LATIN SMALL LETTER K WITH OBLIQUE STROKE;Ll;0;L;;;;;N;;;A7A2;;A7A2
A7A4;LATIN CAPITAL LETTER N WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A5;
A7A5;LATIN SMALL LETTER N WITH OBLIQUE STROKE;Ll;0;L;;;;;N;;;A7A4;;A7A4
A7A6;LATIN CAPITAL LETTER R WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A7;
A7A7;LATIN SMALL LETTER R WITH OBLIQUE STROKE;Ll;0;L;;;;;N;;;A7A6;;A7A6
A7A8;LATIN CAPITAL LETTER S WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A9;
A7A9;LATIN SMALL LETTER S WITH OBLIQUE STROKE;Ll;0;L;;;;;N;;;A7A8;;A7A8
```

Properties to be changed for already encoded letters: (reference to new uppercase letter) 1E9C;LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE;Ll;0;L;;;;in;;;A7A8;;A7A8

## Bibliography:

[1] AlLen, C. G.: A Manual of European Languages for Librarians. London \& New York 1981, ISBN 0-85935-028-2
[2] BeZZenberger, Adalbert: Lettische Dialekt-Studien. Göttingen 1885
[3] Bielenstein, A: Lettische Grammatik. Mitau 1863
[4] FAULMANN, Carl: Das Buch der Schrift. Wien 1880. Reprint Nördlingen 1985, ISBN 392156851X
[5] Hellwig, Wilhelm: Der Satz und die Behandlung fremder Sprachen. 5th. ed., Frankfurt/M. 1930
[6] KAULFÜRST, Fabian (Sorbisches Institut / Serbski Institut, Bautzen/Budyšin, Germany): personal communication 2005/2008/2009.
[7] ReIChSDRUCKEREI: Alphabete und Schriftzeichen des Morgen- und des Abendlandes. Berlin 1924.
[8] WIschmann, Johann: Der Unteutsche Opitz. Oder Kurtze Anleitung zur Lettischen Dicht=Kunst. (The Not-German [i.e. Latvian] Opitz. Or, Short Guidance to the Latvian Art of Poetry.) Rīga 1697
[9] Agenda. sa tu evangelisku zerkwu we Pschuskej. Worjejce (Hoyerswerda; Germany), 1898.
[10] Skohlas-Likkumi. Rīga 1789.

## Acknowledgements:

The proposal was mainly compiled by Karl Pentzlin.
Special thanks for valuable information to:
Fabian Kaulfürst (Sorbisches Institut / Serbski Institut [Sorbian Institute], Bautzen/Budyšin, Germany) Prof. Stefan Kessler (Institut für Fremdsprachliche Philologien / Baltistik der Universität Greifswald, Germany) Gunars Lucans

## Sonja Pentzlin

Andreas Stötzner (Signographisches Institut Andreas Stötzner, Pegau/Sa., Germany) and all members of the Unicode mailing list who participated in the discussion of earlier drafts.

LETTISCH．

| Antiqua | Fraktar | Wert | Antiqua | Fraktur | Wert | Antiqu． | Fraktur | Wert |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| A 2 | $\mathscr{X} \pi$ | a | J j | \＃ | $y$ | Pj pj | Fip $\mathfrak{p}$ | $\mathrm{p}^{\prime}$ |
| $\hat{\mathrm{A}}$ ลิ | $\hat{\mathfrak{1}} \mathbf{a}$ | a | K k | 我 | 1 k hart | R r | ¥ r | r |
| B b | 36 | b | KK | （ ${ }^{2}$ | k lind | （H） | SH0 | $\mathbf{r}^{\prime}$ |
| Bj bj | Bj 1 | $\mathrm{b}^{\prime}$ | L． 1 | Q I | 1 | S f | § i | 2 |
| D I | （2） 1 | d | L． 1 | ¢ | $\mathrm{I}^{\prime}$ | （5） | （3＋3） | 8 |
| Ds ds | W） | d 2 | M m | 3 mm | int | Sch sch | Scif fij | 2 |
| Dsch dsch | Wid）dich | dz | Mj mj |  | mi＇ | Ech fich | （E），fot | s̆ |
| E e | 『8 | e | N $n$ | 9 sin | n | T t | I 1 | t |
| $\stackrel{1}{E}^{\text {e }}$ | （6） | E | ng | mg | ถ． | Teshtuch |  | Ls |
| 6 g | （6） g | g hast． | HF Hr | R \＃ | n＇ | C u | 11 u | u |
| （4．8） | （65） 88 | ¢ Tivul | H20 | （1）${ }^{1}$ | $\mathrm{H}^{\prime}$ | $\hat{\mathrm{t}} \mathrm{u}$ | $\mathfrak{\mathfrak { t }} \hat{\mathrm{H}}$ | ū |
| H H | 94 |  | 00 | $\bigcirc 0$ | 0 | W w | $\mathfrak{W} \mathrm{m}$ | w |
| I i | 31 | i | O\％ | 人 ${ }^{\text {of }}$ | ถ̄ | Wj wj | \％ix tuj | $\mathrm{w}^{\prime}$ |
| İ $\uparrow$ | $\hat{3} \mathrm{t}$ | I | $\mathrm{F}^{1}$ | 98 | p | Z 2 | 83 | d2 |

Fig．1：Scan from［4］，p．231，listing all letters of the Latvian orthography valid at its date of original print （1880），including all letters proposed here（the LATIN SMALL LETTER S WITH OBLIQUE STROKE is found in its round form in the＂Fraktur＂（Blackletter）column）．

## LETTISCH

| $\begin{array}{\|c} \begin{array}{r} \text { Zeiel } \\ \text { alte } \\ \text { Orthogr. } \end{array} \\ \hline \end{array}$ | $\begin{gathered} \text { neue } \\ \text { Orthogr. } \end{gathered}$ | Lautwert | $\begin{array}{\|c\|} \text { alteie } \\ \text { Orthogr. } \\ \hline \end{array}$ | hen Orthogr | Lautwert | Allgemeines．Die lettische Sprache wird in Kurland，im südlichenLivland und in einem kleinen Teil des Gou－ vernements Witebsk gesprochen． Zu ihrer Wiedergabe bedient man sich meist der Fraktur．Seit 1921 ist eine neue Rechtschreibung eingeführt， die die Antiqua verwendet．Sie ist eine rein phonetische and bedient $\operatorname{sich} u$ ．a．der Zeiehen $\check{c}$ ，$s{ }_{s}$ und $\not \approx$ für die entsprechenden Zeichen $t \ddagger f$ ， fiff und $\{d j$ der alten Orthographie． <br> Silbentrennung．Diejenigen Buch－ staben，die einen Laut bilden， dürfen nicht getrennt werden． Ebenso darf $h$（das in der alten Orthographie als Dehnungszeichen dient）von dem vorhergehenden Vokal nicht getrennt werden． <br> Interpunktion．Wie im Deutschen． <br> Literatur． <br> Grammatiken von Bielenstein（1863） und Brentano；Wörterbuch von Utmann u．Brasche． <br> Endselin u．Mühlenbach，Latwenšu gramatika（1907）． |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathfrak{d}$ | a | $a$ | \＃ | 71 | $n$（fast wie ni） |  |
| $\hat{\mathfrak{a}}$ | a | $\bar{a}$ | $1)$ | 0 | 0 |  |
| 6 | b | $b$ | $\hat{\mathrm{D}}$ | $\overline{0}$ | $\bar{o}$ |  |
| 0 | d | $d$ | $p$ | $p$ | $p$ |  |
| e | e | $e$ | $\mathfrak{x}$ | r | $r$ |  |
| $\hat{\mathbf{e}}$ | $\overline{\mathrm{e}}$ | $\vec{e}$ | （1） | r | $r$（fast wie ri） |  |
| $f$ | $f$ | $f$ |  | Z |  |  |
|  | 8 | $g$ | fcti | Ž | $s c h$（stimmhaft） |  |
| （ ${ }^{4}$ | g | $g$（fast wi | F 8 | S |  |  |
| t | h | Hilfszeichen |  |  |  |  |
| i | i | $i$ | 1 | s |  |  |
|  | 1 |  | t | t | $t$ |  |
|  |  |  | u | u | 26 |  |
|  | k |  | $\mathfrak{4}$ | $\overrightarrow{\mathrm{u}}$ | $\bar{u}$ |  |
|  |  |  | 10 | v | $w$ |  |
|  |  |  |  | c | $z(=t \beta)$ |  |
|  |  |  | D）d | dž | $d s c h$（stimmhaft） |  |
|  | $\frac{1}{3}$ | $l$（fast wie b） |  |  |  |  |
| $\mathfrak{m}$ | m | $m$ | － | dz | d） |  |
| $\mathfrak{n}$ | n | $n$ | tF（6） | c | tsch |  |

76
Fig．2：Scan from［7］，p．76，showing the small letterforms and documenting the date until the old Latvian orthography was valid（starting in the sixth line of the right column，translating＂Since 1921，a new orthography is introduced ．．．＂）－Scan and reference provided by Andreas Stötzner，Signographisches Institut

| Schriftzeichen |  | Lautwert | Schriftzeichen |  | Lautwert |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Alte Rechtschreibung | Neue Rechtschreibung |  | AlterRechtschreibung | Neue Rechtschreibung. |  |
| $\mathfrak{a}$ | a | a | (i) | $n$ | $\mathrm{n}^{\mathbf{j}}$ |
| $\mathfrak{a}$ | $\bar{a}$ | $\overline{\text { a }}$ (lang) | (b) | (0) | $\bigcirc$ |
| b | b | b | $\hat{\mathbf{v}}$ | $\overline{0}$ | $\bigcirc$ (lang) |
| d | d | d | $\ddagger$ | p | p . |
| e | e | e | $\mathfrak{r}$ | r | r |
| $\hat{\mathbf{e}}$ | è | $\overline{\mathrm{e}}$ (lang) | ( $)$ | $\frac{1}{8}$ | $\mathrm{r}^{\text {j }}$ |
|  | f |  | 1. | \% | sanftes S |
| , g | g | $g$ | (id) | ż | stimmhaftes sch |
| (3) | $\stackrel{8}{8}$ | g ${ }^{\text {j }}$ |  | S | (wie franz.j) |
| $b$ | h | (nur Hilfs- | fe | S | 13 sch |
| i | 1 | ${ }_{\text {z }} \mathrm{i}$ | t ${ }_{\text {d }}$ | t |  |
| $\hat{\imath}$ | i | $\overline{\mathbf{i}}$ (lang) | 14 | u | u |
| I | j | j | $\hat{\mathfrak{u}}$ | $\overline{\mathrm{u}}$ | $\stackrel{\rightharpoonup}{\mathrm{u}}$ (lang) |
| 1 | k | k | $\mathfrak{w}$ | v | W |
| ( | - ${ }^{\text {k }}$ | $\mathrm{k}^{\mathrm{j}}$ | 8 | c | z, tB |
| I | I | . | Did | dž | dsch (stimmhaft) |
| 1 | 1 | 11 | Dil | dz | df (sanft) |
| $\mathfrak{m}$ | m | m | t $\ddagger$ ( ${ }^{\text {d }}$ | c | tsch |
| $\mathfrak{n}$ | . | n |  |  |  |

Mehrere unserer Laute kommen im Lettischen nicht vor, wie ch ö ü, andere werden umschrieben, z.B. $q=\mathfrak{f w}, x=\mathbf{f}, \mathrm{y}=\mathfrak{i}, \ddot{a}=\mathbf{e}$. Auch h wird als selbständiger Buchstabe nicht gebraucht, sondern nur beim f(f) und als Dehnzeichen bei Selbstlauten, c nur in der neuen Rechtschreibung und als Hilfsbuchstabe bei fid.

## Akzent- und Sonderbuchstaben

2. Die Selbstlaute tragen in der alten Rechtschreibung bei langer Aussprache den Zirkumflex: $\hat{\mathfrak{a}} \hat{\mathbf{e}} \hat{\mathbf{i}} \hat{\mathbf{o}} \hat{\mathfrak{u}}$, mit Trema können $\mathfrak{f}$ und $\mathfrak{i}$ vorkommen, und außerdem gibt es noch mehrere gestrichene Buehstaben:


Fig. 3: Scan from [5], p. 186, showing the small letterforms in the table and the capital forms in the text (last line). - Scan and reference provided by Andreas Stötzner, Signographisches Institut

|  |
| :---: |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Fig. 4: Scan from [8], p.19, showing all lower case forms of the proposed letters.
The Text is in German, discussing rhyming rules for Latvian. The sentence containing the letter series translates: "You also may rhyme the soft [letters] with their hard [counterparts], and the slashed $g, k, l, n, r, s, s c h$ with $g, k, l, n, r, s, s c h$, if only this happens not too often and without necessity."

- Scan and reference provided by Prof. Stefan Kessler, Universität Greifswald

Deem nofault, unt fleatramt butg atbilbeft: Sidhe! Ras weblaf nabe, tam butss pee fawas 2tnafly tharas ar ibfeent un taiqneem $\mathfrak{W a f t s}$ beem nisbildinatees, lapefs miafth ne agrat nabsis.
2. ©toblasbetrunem wiffu jabarta, Po Stobly meifters paivefly, wiana 刃natsiiftianu labbi usellaufibami. Teeveenam butis atbilideft, ne Ia tam, Eam waizalts tolty, un ia fothis ne mabe atbildeft, tad wate tas, fas to finna,

 meffe tam wif's ne buthe runtaft, pirms fa Efofitmeifters par to W3aflu Denvis, un tad jarunna ar (5obsiou.

Fig. 5: Scan from [10], p.12, showing some of the proposed letters.
Retrieved from http://valoda.ailab.Iv/korpuss/senie/faksimili/18gs/SL1789/0012a.htm, 2008-10-28.
Note: This example shows ligatures of two long s with stroke (line starting with " 2 .") and of two I with stroke (second last line), employing a common stroke between the two letters. Such ligatures are common in 18th century prints but we have found no such ligatures in 19th century prints. We regard such ligatures as presentation forms, like other common Blackletter ligatures like "ch".

- Reference provided by Gunars Lucans

Satwijai tàpat fà wifai Sreetwijai wifa wef)= riba bija jagreejif $\mathfrak{u j}$ fazasIatit, tabeby marbubt
 jafafa, fa ta beefgan zenfufees, famas bubfatanaz
Fig. 6: Scan from [1], p.595, showing a sample text of Latvian (set in Fraktur), showing some of the proposed letters.
© tafim bépdyo nowe tomaritivo ja Dolnu Ružyzu
huitamione, ia tym afo to fare pid)ed 27 lĕtami lě̀ma
fublaue a mexjafajo juz taf tu weétojcz we $\mathfrak{B e}$, až buzzo
nabatut ně̌to bềjfo, ab ne mijas dolntymi Şెerbami tež
zfonfi namafali. $\Omega$ tomu io piß
tafi lejdezy lift napizal a we 100 hobalfadi) 1000 erem=

Fig. 7: Scan from a Lower Sorbian text of 1905.
As this text is set in Blackletter, it shows the $s$ with oblique stroke in its long form only.

- Scan provided by Fabian Kaulfürst


## Stuputh.




 hofái bupuajonty a póntajonty pictod Sebut hajote bucle


 stutcit.
2. S(j) ogomójuy a jmiluy ßog, ńcbjaffi Wóp(a), Twója jmilnofcź



 pomógaj nant, $\mathfrak{I y}$ Đog nafdogo ftrowia, humój naş a wódaj nam nafde





Fig. 8: Scan from [9], providing a Lower Sorbian example from 1898, showing also the upper case of the S with oblique stroke. - Scan and reference provided by Fabian Kaulfürst

45 Tó dỉuu un nâkuschu nakti wi\#nsch palizis pì tâs pirmás méitas; ůtrá di̊ná gâsch pio ůtras méitas, kufa Cẻlta istabá dfiwási, un tó sweizinásch. Tâ atbildési tápat ka pirmá: „labdin, wez' Plêscha! Kas tewi te nesa? Mảns wîrs ir stipris,
tewi nůkóus". Plêscha atbilldésch: ${ }^{\text {na }}$, na! es jow nâzu
50 tew nakatá".
Fig. 9: Scan from [2] (p.48), providing a Latvian example showing some of the proposed letters, printed in Antiqua type.

```
Weens fäger's ja` us jakti
zaur }\mp@subsup{}{}{1}\mathrm{ ) äglu mefchinu,
tur sateek wienu meitinu.
nAk meitina, tu smuk' un daila,
kur tawi doemi stäwe"?
#Man' doemi sta' pee tewis
us ägle mefchina".
Kur doama saulit' lazoeti,
tur tumscha nakts atroad,
kur doama siewimi prezät,
tur oetris isprezäje.
```

Fig. 10: Another scan from [2] (p.48), providing a Latvian example showing some of the proposed letters, printed in Antiqua type.

| 8. ${ }^{\text {8 }}$ - ${ }^{\text {a }}$ |  |
| :---: | :---: |
| bridenét, schelten | bridis, Frist |
| griwa, Mündung | grida, Diele |
| ists, eigentlich, ächt | iss, kurz |
| (tills, Keil | Feilas, Plur. tant. Pfand |
| krits, Kreide | kritu, ich falle |

Fig. 11: Scan from [3] (a Latvian grammar in German), showing Latvian examples in Antiqua type.

# PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC $10646{ }^{1}$ <br> <br> Please fill all the sections $A, B$ and $C$ below. 

 <br> <br> 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. <br> 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: Proposal to encode 10 Latin letters for pre-1921 Latvian orthography
2. Requester's name:
3. Requester type (Member body/Liaison/Individual contribution):
4. Submission date:
5. Requester's reference (if applicable):
6. Choose one of the following:
This is a complete proposal:
(or) More information will be provided later:
B. Technical - General
7. 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: |  |

2. Number of characters in proposal: 10
3. Proposed category (select one from below - see section 2.2 of P\&P document):
A-Contemporary B.1-Specialized (small collection)
C-Major extinct $\quad-\cdots .{ }^{\text {Cli. }}$ D-Attested extinct

B.2-Specialized (large collection)
E-Minor extinct

F-Archaic Hieroglyphic or Ideographic
4. Is a repertoire including character names provided? G-Obscure or questionable usage symbols
$\qquad$
b. Are the character shapes attached in a legible form suitable for review?

Yes
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? Karl Pentzlin
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools
used: $\quad$ http://www.pentzlin.com/proposalfont.zip (more information in the text file included in that archive)
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?
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)?
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 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.

[^0]


[^0]:    ${ }^{1}$ Form number: N3152-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05)

