

To: UTC and ISO/IEC JTC1/SC2 WG2

Title: Proposal to encode C WITH BAR

From: Lorna A. Priest (SIL International), Laurentiu Iancu (Microsoft Corporation) and Michael Everson (Everttype)

Date: 14 September 2010

We wish to propose the addition of these characters in the “Latin Extended-D” block. Suggested codepoints are included.

If the characters in this proposal are accepted, the following characters would exist:

€	A79C	LATIN CAPITAL LETTER C WITH BAR • Cambrian symbol
€	A79D	LATIN SMALL LETTER C WITH BAR

## Unicode Character Properties

Properties for the new characters:

A79C;LATIN CAPITAL LETTER C WITH BAR;Lu;0;L;;;;;N;;;A79D;

A79D;LATIN SMALL LETTER C WITH BAR;Ll;0;L;;;;;N;;;A79C;;A79C

There should not be any decompositions.

## Sorting

These characters could be sorted immediately after:

023B;LATIN CAPITAL LETTER C WITH STROKE

023C;LATIN SMALL LETTER C WITH STROKE

## Confusables

These characters could potentially be confused with U+0404 CYRILLIC CAPITAL LETTER UKRAINIAN IE and U+0454 CYRILLIC SMALL LETTER UKRAINIAN IE.

## Usage

Neither of these characters is widely used. LATIN CAPITAL LETTER C WITH BAR was used in Webster’s 1868 dictionary (figure 5). It is also part of the STIX project which was developed “mainly in support of an attempt to gain acceptance into Unicode of additional symbols used in mathematical and technical publishing.”<sup>1</sup> It is in the STIXNonUnicode font (<http://www.stixfonts.org/>) at Private Use Area codepoint U+E0A1. The *Elsevier Science Grid in Unicode* includes the glyph in “Grid B: Symbols” (see figure 6). The Elsevier document also includes a chart listing the position in the grid, the Unicode codepoint, HTML entities and a description of the symbol. This character is listed in the chart without a Unicode mapping and with the description of “Cambrian (era)” (see figure 7). It also appears as a symbol for the Cambrian period in the FGDC *Digital Cartographic Standard for Geologic Map Symbolization* (see figure 9). The example in figure 8 used U+0404 CYRILLIC CAPITAL LETTER UKRAINIAN IE to represent this symbol. It could be encoded as a letterlike symbol (CAMBRIAN SYMBOL). However,

<sup>1</sup> STIX Project Home Page (<http://www.ams.org/STIX/>).

since Pullum and Ladusaw cite use of a LATIN SMALL LETTER C WITH BAR in relation to phonetic transcriptions (see figures 1–3), which was also used in McGuffey’s 1853 Reader (see figure 4), we believe a case pair should be encoded. No examples of them used as a case pair have been seen. Examples of usage are in the Figures section.

## Figures

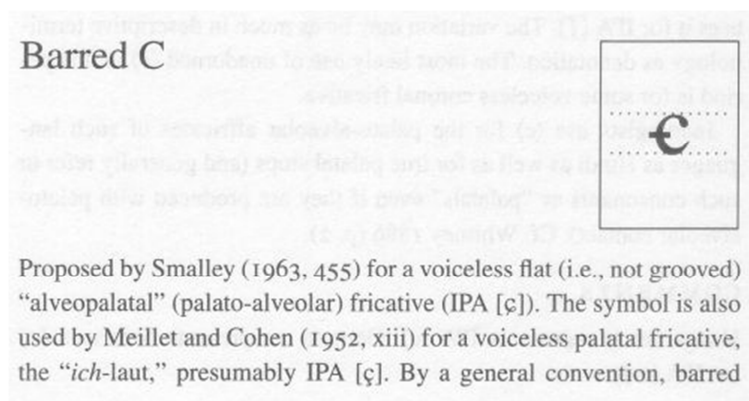


Figure 1. LATIN SMALL LETTER C WITH BAR (Pullum, 1996, p. 28)

Review and Summary 455

Alveolar	Alveopalatal	Retroflexed	Palatal	Mid Velar	Backed Velar	Uvular	Glottal
t	c	ṭ	k̟	k	k̠		ʔ
tʰ	cʰ	ṭʰ	k̟ʰ	kʰ	k̠ʰ		
d	j	ḍ	g̟	g	g̠		
dʰ	jʰ	ḍʰ	g̟ʰ	gʰ	g̠ʰ		
ʈ		(ʈ̣)	(d̟̣)				
ɖ							
(θ)	(ç)		x̟	x	x̠		h
(ð)	(ʝ)		ɣ̟	ɣ	ɣ̠		

Figure 2. LATIN SMALL LETTER C WITH BAR (Smalley, 1963, p. 455)

de l'élément spirant :  $p^f$ , etc. On note de même les consonnes à double occlusion :  $k^p$ .

### Spirantes

Labiales :	sourde : $p$	sonore : $b$
Labio-dentales :	$f$	$v$
Interdentales :	$t$	$d$
Sifflantes :	$s$	$z$
Sifflantes mi-chuintées :	$ś$	$ž$
Chuintantes :	$š$	$ž$
Palatales :	$ç$ (ich-Laut)	
	$x$ (ach-Laut)	
Vélaires :	$k$ ou $h$	$g^1$
Laryngales ou faucales :	$h$	$ε$
Souffle :	$h$	

Figure 3. LATIN SMALL LETTER C WITH BAR (Meillet, 1952, p. xiii).

### TABLE OF ASPIRATES.

$f$ , as in fife.	$t$ , as in tärt.
$h$ , " hīm.	$sh$ , " shē.
$k$ , " <u>eāke.</u>	$ch$ , " chăt.
$p$ , " pīpe.	$th$ , " thīck.
$s$ , " sāme.	$wh$ , " whȳ.

Figure 4. LATIN SMALL LETTER C WITH BAR (McGuffey, 1853, p.6).

VINDI <u>C</u> A-TIVE,	vīn'de-kə-tīv,	vīn-dīk'ə-tīv,
VĪ-O-LON-CEL'LO	} vĕ-o-lon-chēllō, vī-o-lon-sēllō,	
(vĕ-o-lon-chēllō or vĕ-o-lon-sēllō),		
VĪR'TU,	ver-too',	
VIZ'IER (viz'yer),	vīz'yēr,	vīz'yēr,
VOL'UME (vol'yum),	vōl'yūme,	vōl'ūme,
WAIN'S <u>C</u> OT,	wēn'skŭt,	wān'skŏt,
WĀIST <u>C</u> OT,	wēs'kŏt,	wāst'kŏt,
WAR'RIOR (war'-yur),	} war'yur,	war'yur,

Figure 5. LATIN CAPITAL LETTER C WITH BAR (Webster, 1868, p.xx).

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	
a	↔	↘	↓	<	>	†	‡	&	∇	Σ	<	≠	>	≠	≠	≠	≠	≠	×	+		a
b	↔	↘	↓	«	»			¢	∞	Π	≤	≠	≥	≠					x	+		b
c	↔	↘	↓	[	]	§		\$	∞	Π	≤		≥		⊂	⊄	⊃	⊄	·	±	€	c
d	↔	↘	↓	(	)	¶		£	∞	Π	≤	≠	≥	≠	⊂	⊄	⊃	⊄	×	±		d

Figure 6. LATIN CAPITAL LETTER C WITH BAR (Elsevier Science Grid, p. 5).

Position	Unicode	Entities / glyph	Description
Btu	022B4	&ltrie;, &LeftTriangleEqual;, &trianglelefteq;	left triangle, equal
Btv	022B5	&rtrie;, &RightTriangleEqual;, &trianglerighteq;	right triangle, equal
Btw	022EC	&nltrie;, &NotLeftTriangleEqual;, &ntrianglelefteq;	not left triangle, equals
Btx	022ED	&nrttrie;, &NotRightTriangleEqual;, &ntrianglerighteq;	not right triangle, equals
Bu0	02205	&empty;, &emptyv;, &emptyset;, &varnothing;	slashed zero; empty set
Buc	---	<ce:glyph name="camb"/>	Cambrian (era)

Figure 7. Reference to LATIN CAPITAL LETTER C WITH BAR (Elsevier Science Grid, p. 21).

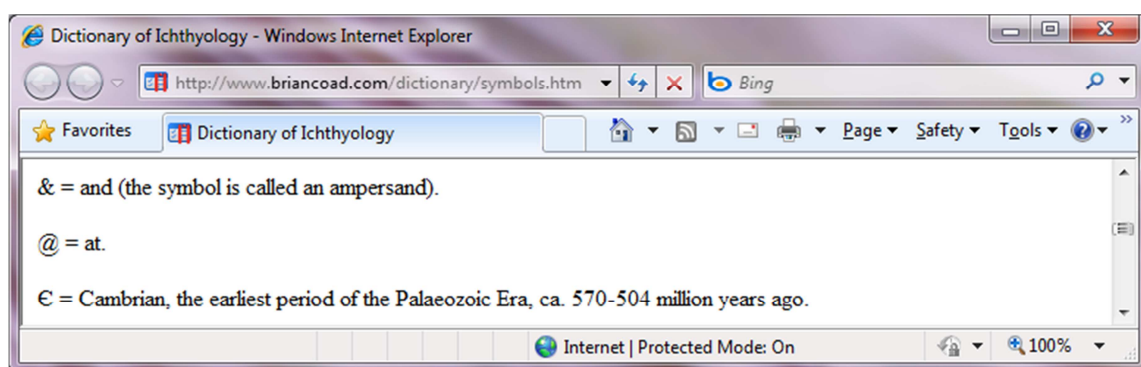


Figure 8. LATIN CAPITAL LETTER C WITH BAR (Coad, 2010).

Cambrian	Period	€
Precambrian	Era	p€

Figure 9. LATIN CAPITAL LETTER C WITH BAR (FGDC, 2010, p. A-32-1).

## References

- Coad, Brian W. and Don E. McAllister. Revised 19 August 2010. . *Dictionary of Ichthyology*.  
<http://www.briancoad.com/dictionary/symbols.htm> (accessed 27 August 2010).
- Federal Geographic Data Committee, ed. 2006. (PDF). FGDC Digital Cartographic Standard for Geologic Map Symbolization FGDC-STD-013-2006. U.S. Geological Survey for the Federal Geographic Data Committee. p. A-32-1. [http://ngmdb.usgs.gov/fgdc\\_gds/geolsymstd/fgdc-geolsym-all.pdf](http://ngmdb.usgs.gov/fgdc_gds/geolsymstd/fgdc-geolsym-all.pdf) (accessed 27 August 2010).
- McGuffey, William Holmes. 1853. *McGuffey's Third Eclectic Reader*. Revised edition. Cincinnati, New York: Wilson, Hinkle & Co.
- Meillet, A., and Marcel Cohen. 1952. *Les Langues du Monde*. Paris: C.N.R.S.
- Pepping, Simon, and Rob Schrauwen. 2004 version. *Elsevier Science Grid in Unicode*.  
[http://www.elsevier.com/framework\\_authors/DTDs/ja50\\_esgrid.pdf](http://www.elsevier.com/framework_authors/DTDs/ja50_esgrid.pdf) (accessed 30 August 2010).
- Pullum, Geoffrey K., and William A. Ladusaw. 1996. *Phonetic Symbol Guide*. Second Edition. Chicago: The University of Chicago Press.
- Smalley, William A. 1963. *Manual of Articulatory Phonetics*. Revised Edition. Tarrytown, NY: Practical Anthropology.
- \_\_\_\_\_. 1989. *Manual of Articulatory Phonetics*. Revised Edition. New York: University Press of America.
- STIX. 2006. STIX Project Home Page (<http://www.ams.org/STIX/>).
- Webster, Noah. 1868. *Webster's Improved Dictionary of the English Language, exhibiting the origin, orthography, pronunciation, & definition of words; embracing all the principal terms used in literature, science & art, according to the best authorities, and likewise giving the synonymous terms for nearly all the words explained*. 2 vols. London, Glasgow, & Edinburgh: William Mackenzie.

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

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:	<b>Proposal to encode C WITH BAR</b>	
2. Requester's name:	<i>Lorna A. Priest (SIL International), Laurentiu Iancu (Microsoft Corporation) and Michael Everson (Everytype)</i>	
3. Requester type (Member body/Liaison/Individual contribution):	Individual	
4. Submission date:		
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:	No	

**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:	Latin Extended-D	
2. Number of characters in proposal:	2	
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?	Lorna Priest	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	SIL International, lorna_priest@sil.org	
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, indexing, transliteration etc. (if yes please enclose information)?	The proposal addresses sorting only	

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.

<sup>2</sup> Form number: N3702-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, 2009-11)

### 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?	Yes
Reference: See proposal	
4. The context of use for the proposed characters (type of use; common or rare)	rare
Reference: See proposal	
5. Are the proposed characters in current use by the user community?	Yes
If YES, where? Reference: See proposal	
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?	
If YES, reference: If possible, should be kept with other related blocks in the BMP	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	Yes
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?	
If YES, reference: See 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?	
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:	