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

Doc Type: **Working Group Document**
Title: **Preliminary proposal for encoding the Palmyrene script in the SMP of the UCS**
Source: **Michael Everson**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2010-01-21**

1. Introduction. The Palmyrene alphabet was used from the first century BCE, in a small independent state established near the Red Sea, north of the Syrian desert between Damascus and the Euphrates. The alphabet was derived as a national script by modification of the customary forms that cursive Aramaic which themselves developed during the first Persian Empire.

Palmyrene is known from documents distributed over a period from the year 9 BCE until 273 CE, the date of the sack of Palmyra by Aurelian. Palmyrene has both a monumental and a cursive form. Fonts for Palmyrene typically follow the monumental style as in the examples below (just as is typical for Latin and Greek). No documents on perishable materials have survived; there are a few painted inscriptions, but many inscriptions on stone. It appears that the monumental forms were derived from the cursive. Earlier inscriptions show more rounded forms, but later inscriptions tended to regularize the letterforms, perhaps in imitation of the monumental the Greek script.

Apparently, cursive script has more ligatures than the monumental; more research is required to determine the nature of ligation in Palmyrene. The letter ALEF, at least, is said to have a final form, and BETH often connects with a following DALETH. At a certain point, some letterforms became confused and in Palmyrene a distinguishing diacritical dot was introduced. This may have been the first appearance of such a feature in the Western Semitic scripts.

2. Processing. Palmyrene is written from right to left horizontally. Palmyrene language inscriptions usually have no space between words; modern editors tend to insert U+0020 SPACE.

3. Character names. The names used for the characters here are based on those used for Imperial Aramaic, Other West Semitic names may have some currency, but the UCS Imperial Aramic names have been preferred here since Palmyrene is an Aramaic language.

4. Numerals. Palmyrene numerals are built up out of 1, 2, 3, 4, 5, 10, 20, and 100. The numbers 2 *וו* and 3 *ווו* and 4 *וווו* are composed of multiples of 1 *ו*, but because in practice the numbers are clumped together as units separate from one another they are encoded as individual characters. The origin of the highest numbers in their Aramaic predecessors is clear; compare Aramaic numbers 10 ➐, 20 ➐ (in origin two 10s one atop the other), and 100 ↩ (in origin two a 10 with a stroke added to differentiate it from 100) with Palmyrene 10 ➝, 20 𝟑, 100 ➝. In the inscriptions number 100 tends to look quite like 10, but they are encoded distinctly here, and the 100 given a slight upturn which occasionally occurs in the inscriptions. The numbers have right-to-left directionality. In the chart below, the third and sixth columns are displayed in visual order.

Value		Value	
1	1 ←	11	1 + 10 ←
2	2 ←	12	2 + 10 ←
3	3 ←	13	3 + 10 ←
4	4 ←	14	4 + 10 ←
5	5 ←	15	5 + 10 ←
6	1 + 5 ←	16	1 + 5 + 10 ←
7	2 + 5 ←	17	2 + 5 + 10 ←
8	3 + 5 ←	18	3 + 5 + 10 ←
9	4 + 5 ←	19	4 + 5 + 10 ←
10	10 ←	100	100 + 1 ←
20	20 ←	200	100 + 2 ←
30	10 + 20 ←	300	100 + 3 ←
40	20 + 20 ←	400	100 + 4 ←
50	10 + 20 + 20 ←	500	100 + 5 ←
60	20 + 20 + 20 ←	600	100 + 1 + 5 ←
70	10 + 20 + 20 + 20 ←	700	100 + 2 + 5 ←
80	20 + 20 + 20 + 20 ←	800	100 + 3 + 5 ←
90	10 + 20 + 20 + 20 + 20 ←	900	100 + 4 + 5 ←

5. Punctuation. Script-specific punctuation for Palmyrene is not known

6. Issues. Two issues need to be resolved before Palmyrene is encoded:

6.1. What is the specific nature of ligation or joining behaviour in Palmyrene? Is it optional or obligatory? What letters does it affect?

6.2. What letter or letters is a diacritical dot used with? Should this (or these) be encoded with a generic dot, with a script-specific dot, or with a precomposed character bearing a dot?

7. Unicode Character Properties

```
10860;PALMYRENE LETTER ALEPH;Lo;0;R;;;;;N;;;;;
10861;PALMYRENE LETTER BETH;Lo;0;R;;;;;N;;;;;
10862;PALMYRENE LETTER GIMEL;Lo;0;R;;;;;N;;;;;
10863;PALMYRENE LETTER DALETH;Lo;0;R;;;;;N;;;;;;
10864;PALMYRENE LETTER HE;Lo;0;R;;;;;N;;;;;
10865;PALMYRENE LETTER WAW;Lo;0;R;;;;;N;;;;;
10866;PALMYRENE LETTER ZAYIN;Lo;0;R;;;;;N;;;;;;
10867;PALMYRENE LETTER HETH;Lo;0;R;;;;;N;;;;;;
10868;PALMYRENE LETTER TETH;Lo;0;R;;;;;N;;;;;;
10869;PALMYRENE LETTER YODH;Lo;0;R;;;;;N;;;;;;
1086A;PALMYRENE LETTER KAPH;Lo;0;R;;;;;N;;;;;;
1086B;PALMYRENE LETTER LAMEDH;Lo;0;R;;;;;N;;;;;;
1086C;PALMYRENE LETTER MEM;Lo;0;R;;;;;N;;;;;
1086D;PALMYRENE LETTER NUN;Lo;0;R;;;;;N;;;;;
1086E;PALMYRENE LETTER SAMEKH;Lo;0;R;;;;;N;;;;;;
1086F;PALMYRENE LETTER AYIN;Lo;0;R;;;;;N;;;;;;
10870;PALMYRENE LETTER PE;Lo;0;R;;;;;N;;;;;
10871;PALMYRENE LETTER SADHE;Lo;0;R;;;;;N;;;;;;
10872;PALMYRENE LETTER QOPH;Lo;0;R;;;;;N;;;;;;
10873;PALMYRENE LETTER RESH;Lo;0;R;;;;;N;;;;;;
10874;PALMYRENE LETTER SHIN;Lo;0;R;;;;;N;;;;;;
10875;PALMYRENE LETTER TAW;Lo;0;R;;;;;N;;;;;
10878;PALMYRENE NUMBER ONE;No;0;R;;;;1;N;;;;;
10879;PALMYRENE NUMBER TWO;No;0;R;;;;2;N;;;;;
1087A;PALMYRENE NUMBER THREE;No;0;R;;;;3;N;;;;;
1087B;PALMYRENE NUMBER FOUR;No;0;R;;;;4;N;;;;;
1087C;PALMYRENE NUMBER FIVE;No;0;R;;;;5;N;;;;;
1087D;PALMYRENE NUMBER TEN;No;0;R;;;;10;N;;;;;
1087E;PALMYRENE NUMBER TWENTY;No;0;R;;;;20;N;;;;;
1087F;PALMYRENE NUMBER ONE HUNDRED;No;0;R;;;;100;N;;;;;
```

8. Bibliography

Ballhorn, Friedrich. 1864. *Alphabete orientalischer und occidentalischer Sprachen*. Leipzig: Brockhaus.

Cantineau, J. 1935. *Grammaire du palmyrénien épigraphique*. Le Caire: Imprimerie de l'Institut français d'archéologie orientale, 1935.

Christian, Arthur. 1905. *Débuts de l'imprimerie en France: L'Imprimerie Nationale; L'Hôtel de Rohan*. Paris: Imprimerie Nationale.

Faulmann, Carl. 1990 (1880). *Das Buch der Schrift*. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8

Fossey, Charles. 1948. *Notices sur les caractères étrangers anciens et modernes rédigées par un groupe e savants*. Nouvelle édition míse à jour à l'occasion du 21e Congrès des Orientalistes. Paris: Imprimerie Nationale de France.

Healey, John F. 1990. *The early alphabet*. (Reading the past). London: British Museum. ISBN 0-7141-8073-4

Ifrah, Georges. 1998. *The universal history of numbers: from prehistory to the invention of the computer.* London: Harvill Press. ISBN 1-86046-324-X

Imprimerie Nationale. 1990. *Les caractères de l'Imprimerie Nationale*. Paris: Imprimerie Nationale Éditions. ISBN 2-11-081087-8

Lidzbarski, Mark. 1962 (1898). *Handbuch der nordsemitischen Epigraphik nebst ausgewählten Inschriften*. Hildesheim: Georg Olms Verlagsbuchhandlung.

Naveh, Joseph. 1987. *Early history of the alphabet: an introduction to West Semitic epigraphy and palaeography*. Jerusalem: Magnes Press, the Hebrew University. ISBN 965-223-436-2

	1086	1087
0	𐡠 10860	𐡰 10870
1	𐡡 10861	𐡱 10871
2	𐡢 10862	𐡲 10872
3	𐡣 10863	𐡳 10873
4	𐡤 10864	𐡴 10874
5	𐡥 10865	𐡵 10875
6	𐡦 10866	
7	𐡧 10867	
8	𐡨 10868	𐡸 10878
9	𐡩 10869	𐡹 10879
A	𐡪 1086A	𐡺 1087A
B	𐡫 1086B	𐡻 1087B
C	𐡬 1086C	𐡼 1087C
D	𐡭 1086D	𐡽 1087D
E	𐡮 1086E	𐡾 1087E
F	𐡯 1086F	𐡿 1087F

Letters

10860	𐡠	PALMYRENE LETTER ALEPH
10861	𐡡	PALMYRENE LETTER BETH
10862	𐡢	PALMYRENE LETTER GIMEL
10863	𐡣	PALMYRENE LETTER DALETH
10864	𐡤	PALMYRENE LETTER HE
10865	𐡥	PALMYRENE LETTER WAW
10866	𐡦	PALMYRENE LETTER ZAYN
10867	𐡧	PALMYRENE LETTER HETH
10868	𐡨	PALMYRENE LETTER TETH
10869	𐡩	PALMYRENE LETTER YODH
1086A	𐡪	PALMYRENE LETTER KAPH
1086B	𐡫	PALMYRENE LETTER LAMEDH
1086C	𐡬	PALMYRENE LETTER MEM
1086D	𐡭	PALMYRENE LETTER NUN
1086E	𐡮	PALMYRENE LETTER SAMEKH
1086F	𐡯	PALMYRENE LETTER AYN
10870	𐡰	PALMYRENE LETTER FE
10871	𐡱	PALMYRENE LETTER SADHE
10872	𐡲	PALMYRENE LETTER QOPH
10873	𐡳	PALMYRENE LETTER RESH
10874	𐡴	PALMYRENE LETTER SHIN
10875	𐡵	PALMYRENE LETTER TAW

Numbers

10878	𐡸	PALMYRENE NUMBER ONE
10879	𐡹	PALMYRENE NUMBER TWO
1087A	𐡺	PALMYRENE NUMBER THREE
1087B	𐡻	PALMYRENE NUMBER FOUR
1087C	𐡼	PALMYRENE NUMBER FIVE
1087D	𐡽	PALMYRENE NUMBER TEN
1087E	𐡾	PALMYRENE NUMBER TWENTY
1087F	𐡿	PALMYRENE NUMBER ONE HUNDRED

10. Figures.

Figure 1. Chart of Semitic alphabets from Ballhorn 1864, showing Palmyrene.

SCHRIFTEN IM OSTEN PALÄSTINA'S.

Hauranitisch		Nabathäisch	Palmyrenisch	Wert
nach HALÉVY	H. MÜLLER			
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	a
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	b
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	g
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	d
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	h
(Zeichen)		*(Zeichen)*	*(Zeichen)*	w
(Zeichen)		*(Zeichen)*	*(Zeichen)*	z
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	χ
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	ṭ
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	y
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	k
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	l
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	m
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	n
(Zeichen)		*(Zeichen)*	*(Zeichen)*	s
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	a
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	p
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	ṣ
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	q
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	r
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	š
(Zeichen)	*(Zeichen)*	*(Zeichen)*	*(Zeichen)*	t

In der Wüste Hauran, welche sich zwischen Palästina und Arabien erstreckt, befinden sich Ruinen mit merkwürdigen Inschriften, welche ein Mittelglied zwischen himyarischen und aramäischen Formen bilden. An sie schliessen sich die palmyrenische und die Schrift der Nabathäer, deren Hauptstadt Petra war; die hier gegebenen Zeichen sind theils den Münzen aus dem 2. Jahrhundert v. Chr., theils den sinaitischen Inschriften entnommen, deren Entzifferung besonders den Studien LEVY's zu verdanken ist.

Figure 2. Chart of "scripts in the east of Palestine" from Faulmann 1880, showing Palmyrene.

ALPHABETS ARAMÉENS COMPARÉS.

(ARAMÉEN ANCIEN, PALMYRÉNIEN, NABATÉEN.)

Figure 3. Chart of "Aramaic alphabets" from Ballhorn 1864, showing Palmyrene.

TEXTE PALMYRÉNIEN.

DÉDICACE DE L'AN 242-243 DE NOTRE ÈRE,

DOUZE ANS APRÈS LA VENUE D'ALEXANDRE SÉVÈRE À PALMYRE.

mwšn	*wklm*	*rb*	*wklm*	*rb*	*aldbz*	*sylrwa*	*sylwy*	*mlṣ*
Nassoum,	de Malikou	fils,	de Malikou	fils	,Zabdila	Aurelius	de Julius	Statue

swrdnskla	*ahla*	*yd*	*atywtymb*	*aynlkl*	*gṭrṣa*	*awh*	*yd*
Alexandre	dieu	du	lors de la venue	,de la colonie	stratège	était	qui

ydkw	*anwmgyh*	*swnypsrk*	*nnt*	*awh*	*ydk*	*šmšw*	*rsk*
et quand,	le gouverneur	Crispinus	ici	fut	lorsque	Et il a servi	.César

yd	*kwš*	*br*	*awhw*	*naygs*	*nwbz*	*aynygl*	*ty*	*akl*	*yta*
Ce que	.du marché	chef	et il était	,long	un temps	,les légions		ici	il conduisit

4+10+20+20+(100×5)	*tnṣ*	*hrkyl*	*swmdw*	*alwb*	*hl*	*myka*
.554	l'an	en son honneur	et le peuple	le sénat	à lui	ont élevé

Figure 4. Sample text in Palmyrene from Fossey 1948.

Figure 4. Sample of Palmyrene text from Christian 1905.

A. Administrative

1. Title
Preliminary proposal for encoding the Palmyrene script in the SMP of the UCS
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2010-01-21
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
No.
6b. More information will be provided later
Yes.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.
1b. Proposed name of script
Palmyrene.
1c. The proposal is for addition of character(s) to an existing block
No.
1d. Name of the existing block
2. Number of characters in proposal
30.
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)
Category E.
4a. Is a repertoire including character names provided?
Yes.
4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
4c. Are the character shapes attached in a legible form suitable for review?
Yes.
5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson.
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, FontLab.
6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes.
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
7. 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.
8. 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/Public/UNIDATA/ UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.
See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
No.
2a. 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.
2b. If YES, with whom?
Reinhard Lehmann, John Healey, and Viktor Golinets.

2c. 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?

See above.

4a. The context of use for the proposed characters (type of use; common or rare)

To write the Aramaic language.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

In scholarly publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?