

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

Doc Type: Working Group Document**Title: Revised proposal to encode the Manichaean script in the UCS****Source: Michael Everson****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2002-12-03****A. Administrative****1. Title**

Revised proposal to encode the Manichaean script in the UCS.

2. Requester's name

Michael Everson

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2002-12-03

5. Requester's reference (if applicable)**6. Choose one of the following:****6a. This is a complete proposal**

Yes.

6b. More information will be provided later

No.

B. Technical -- General**1. Choose one of the following:****1a. This proposal is for a new script (set of characters)**

Yes.

Proposed name of script

Manichaean.

1b. The proposal is for addition of character(s) to an existing block

No.

1b. Name of the existing block**2. Number of characters in proposal**

36

3. Proposed category (see section II, Character Categories)

Category B.1.

4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)

Level 3.

4b. Is a rationale provided for the choice?

Uses combining characters.

4c. If YES, reference**5a. Is a repertoire including character names provided?**

Yes.

5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

Yes.

5c. Are the character shapes attached in a legible form suitable for review?

Yes.

6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No.

7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

No.

8. 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. 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.

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?

Desmond Durkin-Meisterernst and the TITUS project in Frankfurt.

2c. If YES, available relevant documents

<http://titus.fkidg1.uni-frankfurt.de/unicode/iranian/3tagung.htm>

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Iranianists and other scholars.

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

Uncommon; the script is important for students of the Manichaean religion, as well as Persian, Parthian, Sogdian, Bactrian, Turkish, and Tokharian.

4b. Reference

P. Oktor Skjærvø in Daniels & Bright 1996.

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

Yes.

5b. If YES, where?

Scholarly publications.

6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?

No. Proposed allocation is U+10B00-U+10B2F

6b. If YES, is a rationale provided?

6c. If YES, reference

Accordance with the Roadmap.

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

12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

12b. If YES, reference

13a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

13b. If YES, describe in detail (include attachment if necessary)

14a. Does the proposal contain any Ideographic compatibility character(s)?

No.

14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

The Manichaean script evolved from Syriac Estrangelo, and is believed by some to have been devised by Mani himself, who founded the Manichaean religion in the third century C.E. The script is written from right to left. A few of the characters have final forms in addition to the form which is used initially and medially. A set of punctuation marks exist, and there are four diacritical marks which, while they look similar, do behave differently. The double dot used to differentiate consonants is the same size as the single dot used for a similar purpose.

TABLE XXX - Row 1xx: MANICHAEAN

	1xx0	1xx1	1xx2
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
A			
B			
C			
D			
E			
F			

G = 00
P = 01

TABLE XXX - Row 1xx: MANICHAEAN

hex	Name	hex	Name
00	MANICHAEAN PUNCTUATION FLEURON		
01	MANICHAEAN PUNCTUATION DOT WITHIN DOT		
02	MANICHAEAN PUNCTUATION SINGLE THICK DOT		
03	MANICHAEAN PUNCTUATION SINGLE THIN DOT		
04	MANICHAEAN PUNCTUATION TWO DOTS		
05	MANICHAEAN PUNCTUATION LINE FILLER		
06	(This position shall not be used)		
07	(This position shall not be used)		
08	(This position shall not be used)		
09	(This position shall not be used)		
0A	(This position shall not be used)		
0B	(This position shall not be used)		
0C	(This position shall not be used)		
0D	(This position shall not be used)		
0E	(This position shall not be used)		
0F	(This position shall not be used)		
10	MANICHAEAN LETTER ALAPH		
11	MANICHAEAN LETTER BETH		
12	MANICHAEAN LETTER GAMAL		
13	MANICHAEAN LETTER GHAMAL		
14	MANICHAEAN LETTER DALATH		
15	MANICHAEAN LETTER HE		
16	MANICHAEAN LETTER WAW		
17	MANICHAEAN LETTER ZAIN		
18	MANICHAEAN LETTER HETH		
19	MANICHAEAN LETTER TETH		
1A	MANICHAEAN LETTER YUDH		
1B	MANICHAEAN LETTER KAPH		
1C	MANICHAEAN LETTER LAMADH		
1D	MANICHAEAN LETTER DHALATH		
1E	MANICHAEAN LETTER MIM		
1F	MANICHAEAN LETTER NUN		
20	MANICHAEAN LETTER SEMKATH		
21	MANICHAEAN LETTER AYIN		
22	MANICHAEAN LETTER PE		
23	MANICHAEAN LETTER SADHE		
24	MANICHAEAN LETTER JAIN		
25	MANICHAEAN LETTER QAPH		
26	MANICHAEAN LETTER RISH		
27	MANICHAEAN LETTER SHIN		
28	MANICHAEAN LETTER TAW		
29	MANICHAEAN COMBINING SINGLE DOT		
2A	MANICHAEAN COMBINING DOUBLE DOT		
2B	MANICHAEAN COMBINING ABBREVIATION DOTS		
2C	MANICHAEAN COMBINING PLURAL DOTS		
2D	MANICHAEAN LENGTHENER		
2E	(This position shall not be used)		
2F	(This position shall not be used)		