

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

A. Administrative

1. Title

Proposal to encode two Arabic characters to the UCS

2. Requester's name

Unicode Consortium

3. Requester type

Liaison contribution

4. Submission date

2001-04-04

5. Requester's reference

6a. Completion

This is a complete proposal.

6b. More information to be provided?

No.

B. Technical -- General

1a. New script? Name?

No.

1b. Addition of characters to existing block? Name?

Yes. Arabic

2. Number of characters

2.

3. Proposed category

Category A

4. Proposed level of implementation and rationale

Level 1 because they are non-combining.

5a. Character names included in proposal?

Yes

5b. Character names in accordance with guidelines?

Yes

5c. Character shapes reviewable?

Yes

6a. Who will provide computerized font?

Michael Everson, Everson Gunn Teoranta

6b. Font currently available?

Yes

6c. Font format?

TrueType

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

No.

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

No.

8. Does the proposal address other aspects of character data processing?

No.

C. Technical -- Justification**1. Contact with the user community?**

Yes. Kamal Monsour, Monotype Corporation.

2. Information on the user community?

Scholars.

3a. The context of use for the proposed characters?

To write ancient Arabic text.

3b. Reference**4a. Proposed characters in current use?**

Yes.

4b. Where?

Scholarly contexts.

5a. Characters should be encoded entirely in BMP?

Yes.

5b. Rationale

Should be kept with other Arabic characters.

6. Should characters be kept in a continuous range?

No.

7a. Can the characters be considered a presentation form of an existing character or character sequence?

No.

7b. Where?**7c. Reference****8a. Can any of the characters be considered to be similar (in appearance or function) to an existing character?**

No.

8b. Where?**8c. Reference****9a. Combining characters or use of composite sequences included?**

No.

9b. List of composite sequences and their corresponding glyph images provided?

No.

10. Characters with any special properties such as control function, etc. included?

No.

Proposal

This proposal specifies the characters necessary for the writing of Ancient Arabic text. In the early states of Arabic writing, many letters were ambivalent or even multivalent; for instance, a DAL (U+062F) also stood for THAL (U+0630) which bears a dot in Modern Arabic.

Since the alphabet of Modern Arabic consists of many letter shapes which appear in both dotless and dotted forms, in most cases there exists already a dotless form of every character. As exceptions to this pattern, the following 2 letterforms are missing from the standard alphabet of Modern Arabic:



ARABIC LETTER DOTLESS BEH 066E

Usage: this letter was used to represent BEH, TEH, THEH, and NUN without distinction.



ARABIC LETTER DOTLESS QAF 066F

Usage: this letter was used the final and separate forms of QAF. Its initial and medial forms are indistinguishable from that of dotless FEH.