2004-05-03

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 add QAMATS QATAN to the BMP of the UCS

Source: Michael Everson & Mark Shoulson

**Status: Individual Contribution** 

**Action:** For consideration by JTC1/SC2/WG2 and UTC

Date: 2004-05-03

This document requests an additional character to be added to the UCS and contains the proposal summary form.

#### A. Administrative

1. Title

Proposal to add QAMATS QATAN to the BMP of the UCS.

2. Requester's name

Michael Everson & Mark Shoulson

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

Individual contribution.

4. Submission date

2004-05-03

- 5. Requester's reference (if applicable)
- 6. Choose one of the following:

6a. This is a complete proposal

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)

Proposed name of script

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

1b. Name of the existing block

Hebrew

2. Number of characters in proposal

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?

4c. If YES, reference

Combining character.

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. TrueType.

**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?

Yes, see bibliography below.

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

Yes.

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)?

Yes, see below.

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.

Yes, see Unicode properties below.

#### 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.)?

No.

2b. If YES, with whom?

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?

Yes.

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

Liturgical typesetting.

4b. Reference

See examples below.

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

Yes.

5b. If YES, where?

See examples below.

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?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

All Hebrew points are in the BMP.

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

N/A

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?

Yes.

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

Yes.

10c. If YES, reference

It is derived from HEBREW POINT QAMATS but it has a different function and shape.

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)?

Yes.

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

No.

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?

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?

14c. If YES, reference

# D. Proposal

One of the Hebrew vowel signs, HEBREW POINT QAMATS has different pronunciations under certain phonological circumstances. A number of important publishers of religious books have been printing this vowel with a distinctive glyph in order to assist their readers in pronouncing pointed Hebrew text correctly. This innovation has produced a new character, identified with new name, HEBREW POINT QAMATS QATAN, or 'short qamats'. Strictly speaking, *qamats qatan* is a grammatical term referring to the changed pronunciation of *qamats* [a]/[ɔ] to [o].

Although these publishers seem to have been working independently, most of them agree on the representation of the QAMATS QATAN. It is usually presented with a longer vertical; it can also be represented simply with a larger glyph, and in one case we have found the horizontal lengthened and bent. In any case, the modified glyph is used consistently in these texts to represent what all agree is QAMATS QATAN.

05BA



HEBREW POINT QAMATS QATAN (glyph variants  $\bigcirc$ ,  $\bigcirc$ )

#### **Unicode Character Properties**

05A2; HEBREW POINT QAMATS QATAN; Mn; 220; NSM;;;;; N;; \*;;;

# Bibliography

Koren Jerusalem Inc. 1995. חמישה חומשי תורה. Jerusalem: Koren.

Riahi, S. M., ed. 2001. "סימנים" ספר תהלים. Jerusalem: Feldheim Press.

Shiovitz, Jeffrey, ed. 2001. B'kol Echad קול אחד. New York: United Synagogue of Conservative Judaism.

Tal, Shlomo, ed. 1976. כדור רנת ישראל: נוסח אשכנו, Jerusalem: Keter Press.

# **Figures**

We will begin with some examples of modern texts which distinguish the two points.

שִׁים שָׁלוֹם טוֹבָה וּבְרָכָה, חֵן נָחֶסֶד וְרַחֲמִים עָלֵינוּ וְעַלֹּכְּל יִשְּׁרָאֵל עַמֶּךָ, בְּרְכֵנוּ אָבִינוּ כְּלָנוּ כְּאָחָד בְּאוֹר פָּנֶידָ, כִּי בְאוֹר פָּנֶידָ נְתַתָּ לְנוּ יי אֱלֹהֵינוּ תּוֹרַת חַיִּים וְאַהֲבַת חֶסֶד, וּצְדָקָה וּבְרָכָה וְרַחֲמִים וְחַיִּים וְשָׁלוֹם. וְטוֹב בְּעֵינֶיךְ לְבָרֵךְ אָת עַמְּךָ יִשְּׂרָאֵל בְּכָל עֵת וּבְכָל שָׁעָה בִּשְּׁלוֹמֶךְ.

**Figure 1.** From the Koren *Torah Haftarot and Sabbath Prayers*. In this picture and those that follow, instances of U+05AA HEBREW POINT QAMATS pointed to with a plain arrow are shown alongside HEBREW POINT QAMATS QATAN pointed to with a feathered arrow.



Figure 2. From Riahi 2001 (the "Simanim" Psalms).

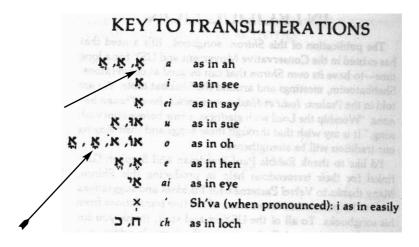


Figure 3. From Shiovitz 2001, a songbook often presented at bar mitzvahs and the like.

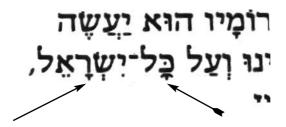


Figure 4. From Shiovitz 2001.

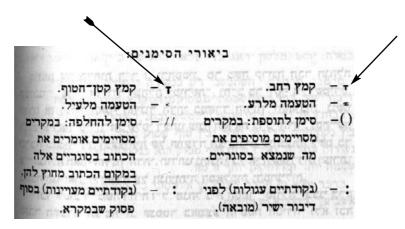
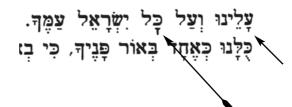


Figure 6. From Tal 1976.



**Figure 7.** From Tal 1976.