

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

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:** Proposal to Encode WORD SEPARATOR MIDDLE DOT

2. Requester's name: US

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

4. Submission date: 2007-09-18

5. Requester's reference (if applicable): L2/07-225 (consensus 112-C23)

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: Supplemental Punctuation (2E00–2E7F)

2. Number of characters in proposal: 1

3. Proposed category (select one from below - see section 2.2 of P&P document):

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 _____

4. Proposed Level of Implementation (1, 2 or 3) (see Annex K in P&P document): 1

Is a rationale provided for the choice? No

If Yes, reference: _____

5. 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? No

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

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

7. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? No

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? No

8. Special encoding issues:

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, suggested character properties are included (see section D)

9. 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.

¹ Form number: N2652-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain _____	<u>No</u>
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.)? If YES, with whom? <u>Paleographers</u> If YES, available relevant documents: _____	<u>Yes</u>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: _____	<u>No</u>
4. The context of use for the proposed characters (type of use; common or rare) Reference: <u>This word-separator punctuation is used in many different historic scripts as well as some modern scripts.</u>	<u>Common</u>
5. Are the proposed characters in current use by the user community? If YES, where? Reference: <u>See comments in Section E</u>	<u>Yes</u>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference: <u>Should be in existing punctuation block</u>	<u>Yes</u>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<u>N/A</u>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference: _____	<u>No</u>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference: _____	<u>No</u>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference: <u>See comments in Section E</u>	<u>Yes</u> <u>Yes</u>
11. Does the proposal include use of combining characters and/or use of composite sequences? 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: _____	<u>No</u>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) _____	<u>No</u>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference: _____	<u>No</u>

D. Proposed Characters

D.1 Character Repertoires

One character is proposed:

U+2E37 WORD SEPARATOR MIDDLE DOT

D.2. Unicode Character Properties

This character should have the similar properties to U+10101 AEGEAN WORD SEPARATOR DOT. In particular, the distinctive properties are:

General Category = Po

Canonical combining class = 0

Bidi category = ON

Line breaking = BA

Script = Common

Pattern_Syntax = True

E. Other Information

A number of dot punctuation characters already exist in the UCS. Multiplication of script-specific middle dots that are minimally distinct in appearance does not serve users well in the long run. Common (non-script-specific) word separators can meet the needs of users and avoid ambiguities. Because the word separator dots are used in scripts of either right-to-left, left-to-right or variable directionality, there is the possibility that users may want to use characters with particular bidirectional properties. Existing characters provide common (non-script-specific) word separators with two of the three relevant bidirectional-category properties:

16EB Runic Word Separator Middle Dot (left-to-right)

1091F Phoenician Word Separator Middle Dot (right-to-left)

(Note: U+16EB has “Runic” in its name, and U+1091F has “Phoenician” in its name. Nonetheless, the script property for these characters is common.) The proposed character fills the gap, providing a character with a bidirectional category of ON (other neutral).