

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 encode a Middle Asterisk in the UCS

Author: Karl Pentzlin

Status: Individual Contribution

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

Date: 2017-05-04

1. Introduction

The symbols universally present on telephone keyboards besides the decimal digits are the number sign and the asterisk, as standardized in the recommendation ITU-T E.161 “Arrangement of digits, letters and symbols on telephones and other devices that can be used for gaining access to a telephone network”, where the symbols are named “square” and “star”, respectively.

While that recommendation specifies exact shapes for these symbols, modern telephone keyboard designs show a wide spectrum of glyph forms for the star, comparable to the asterisk in the wide spectrum of fonts available for printing.

However, the star commonly matches the other symbols (especially the number sign) in size and vertical position (see the examples below).



Fig. 1: Telephone keyboards designed 1982 and 2013, showing different shapes of the star

The Unicode character U+002A ASTERISK, on the other hand, is usually smaller and elevated. In fact, it is a “superscript asterisk” (suitable for footnote references, birth date marking, and for marking reconstructed or ungrammatical forms in linguistics), as shown in the sample collection:

|* I* |* I* I* I* |* I* I* I* I*

Fig. 2: U+002A ASTERISK together with an uppercase I (to mark the caps height) using the fonts: Arial, Times New Roman, Calibri, Cambria, Adobe Garamond Pro, Palatino Linotype, DejaVu Sans, DejaVu Serif, Cardo, Andron Mega Corpus, Courier New, Consolas.

This collection also shows the glyphic variation regarding the basic shape (true star with peaks / spoked asterisk / teardrop asterisk) and the peak/spoke count (five / six).

At this time, Unicode contains no asterisk character which is suited to represent the telephony star in running text (by matching the design and the vertical position of the number sign), allowing the glyph variation of U+002A as well as the one of real existing stars of telephony keyboards.

► Therefore, a new character MIDDLE ASTERISK is proposed here for this purpose.

(The proposed name indicates its vertical positioning similar to U+00B7 MIDDLE DOT.)

The proposed MIDDLE ASTERISK is needed for descriptions of the telephone keyboard.

Die Taste * dient der Steuerfunktion »Befehl eingeben«, die Taste # gibt den Befehl an die Zentrale »Information abrufen«. Das

Fig. 3: MIDDLE ASTERISK in a book describing the German "Bildschirmtext" service, 1983 [3].

Anrufschutz - Gruppe *28* Gruppennr. # #28* Gruppennr. #
Durchbrechen *60* Nebenstelle #

Fig. 4: MIDDLE ASTERISK in an instruction for use of a phone system, 1999 [4].

- The German NB responsible for keyboard layouts is currently evaluating a layout which covers the needs of quality typographic designs. As it will distinguish U+2212 MINUS SIGN from U+2013 EN DASH, it also requires to distinguish the "telephony star" from U+002A ASTERISK. Therefore, the MIDDLE ASTERISK to represent the "telephony star" is urgent.

While there are several asterisks in Unicode, all of these are too specific to allow the glyphic variation of U+002A ASTERISK and the proposed MIDDLE ASTERISK.

- U+2055 FLOWER PUNCTUATION MARK has to match with the Indic scripts it is designed for (usually it has eight spokes).
- U+2217 ASTERISK OPERATOR has to contrast to U+22C6 STAR OPERATOR and therefore is not able to be represented by star-like glyphs (with true peaks rather than spokes). Also, it is constrained by design requests for mathematical operators.
- The asterisks in the "Dingbat" (2700...27BF) and "Geometric Shapes Extended" (1F780...1F7FF) blocks have specific glyph appearances denoted by their name.

Adding the MIDDLE ASTERISK to Unicode does not implicate any special challenge to font designers. The glyph of U+002A ASTERISK can be reused by only adjusting its vertical position (while providing a somewhat enlarged glyph is appropriate).

2. Encoding Considerations

The character is proposed as punctuation mark in the block "Supplemental Punctuation", to keep it in line with the other asterisks which are subject to similar glyphic variation, which all are encoded as punctuation marks:

U+204E LOW ASTERISK
U+2042 ASTERISM
U+2051 TWO ASTERISKS ALIGNED VERTICALLY.

Alternatively, it can be encoded as U+2BFF in the block "miscellaneous symbols and arrows".

Both code points are the first free code points in the respective blocks according to the Pipeline Table on the Unicode website from 2017-01-27, which seems to be current at 2017-05-04.

3. Proposed Character

* U+2E4F MIDDLE ASTERISK
→ 002A asterisk
→ 00B7 middle dot
→ 204E low asterisk
→ 2217 asterisk operator
→ 273D heavy teardrop-spoked asterisk

Properties:

2E4F;LOW ASTERISK;Po;0;ON;;;;;N;;;;;

Alternatively:

2BFF;LOW ASTERISK;So;0;ON;;;;;N;;;;;

4. References

- [1] Detail of the German telephone FeTAp 751, produced 1982 by Siemens for the Deutsche Bundespost (German Federal Mail). Source: <https://commons.wikimedia.org/wiki/File:Detail-Tastatur-FeTAp-751-1982.JPG> – © CC-BY-SA 3.0 by Wikipedia user Rotkaeppchen68
- [2] Detail of Nokia 108 Dual SIM mobil telephone Compact (released Oct 2013). Source: <https://www.amazon.de/Nokia-108-Dual-SIM-Mobiltelefon-Compact/dp/B00GIF5ZQK>
- [3] Hans-Peter Förster, *Bildschirmtext*, Humboldt-Taschenbuchverlag München (Munich, Germany) 1983, ISBN 3-581-66457-7, p. 104
- [4] *Bedienungsanleitung Dialog 3210, 3211 und 3212 Systemtelefone* (Instructions for use for telephone models Dialog 3210, 3211, 3212; German), DE/LZT S 102 2552 R A, Ericsson GmbH 1999, p. 89

**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 a Middle Asterisk in the UCS

2. Requester's name: Karl Pentzlin

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

4. Submission date: 2017-05-04

5. Requester's reference (if applicable): _____

6. Choose one of the following:

This is a complete proposal: Yes

(or) More information will be provided later:

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

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. 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? Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? n/a (the proposed symbol is a positional variant in a row of already encoded characters)

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): _____

6. References:

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

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

7. 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)? no

8. 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 Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3902-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain		No
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? If YES, available relevant documents:	<i>German NB responsible for keyboards</i> <i>see text</i>	Yes
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>All users of telephone keyboards</i>	Yes
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>All users of telephone keyboards</i> <i>see text</i>	
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>see text</i>	Yes
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:	<i>To keep them in line with similar characters</i>	Yes Yes
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		n/a
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:		No
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:		No
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:		No
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:		
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)		
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		No