

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 one Bopomofo character in the UCS****Source: Michael Everson (Yè Mìháó), H. W. Ho, Andrew West (Wèi Ān)****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2006-10-19**

The UCS contains a number of Bopomofo characters in the Bopomofo block (U+3100-U+312F) and the Bopomofo Extended block (U+31A0-U+31BF). One character is missing, and is proposed here for encoding at U+312D, as БОПОМОФО LETTER IH.

The character, 𠂇 (an upside-down U+3113 𠂇 БОПОМОФО LETTER ZH), was once used in the Republic of China to denote the vowel [ə] *i* as in 𠂇 *zhi*, 𠂇 *chi*, 𠂇 *shi*, and 𠂇 *ri* (Wade-Giles spelling *chih*, *ch'ih*, *shih*, and *jih*), and the vowel [ɨ] in 𠂇 *zi*, 𠂇 *ci*, and 𠂇 *si* (Wade-Giles spelling *tzǔ*, *tz'ǔ*, and *szǔ*) when either sound is written in isolation. The form 𠂇𠂇 would be redundant, however, as the [ə] is inherent in the 𠂇. The character was never encoded in standards of either the People's Republic of China or of Taiwan.

Evidence for this character is given in a table of simplified Chinese characters (第一批简体字表) issued by the government of the Republic of China in 1935. The authority for the publication for the table of simplified characters was the “Ministry of Education” (中華民國教育部); it was published 1935-08-21. The tables were reprinted in the book “现行汉字规范问题” (*Xiànxíng Hànzì guīfàn wèntí*, ISBN 7-100-03652-6) published in 2002.

Unicode character properties

312D;BOPOMOFO LETTER IH;Lo;0;L;;;;;N;;;;;

A. Administrative

1. Title

Proposal to encode one Bopomofo character in the UCS

2. Requester's name

Michael Everson (Yè Miháo), H. W. Ho, Andrew West (Wèi Ān)

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

Individual contribution.

4. Submission date

2006-10-19

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)

No.

1b. Proposed name of script

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

Yes.

1d. Name of the existing block

Bopomofo.

2. Number of characters in proposal

1.

3. Proposed category (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)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. 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. 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.

See above.

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?

Linguists.

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

Phonetics.

4b. Reference

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

No.

5b. If YES, where?

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Keep with Bopomofo.

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

No.

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

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?