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

Doc Type: **Working Group Document**
Title: **Evidence for the glyph form of U+17F03**
Source: **Andrew West**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2025-06-05**

This document summarizes the evidence for the glyph form of Tangut ideograph U+17F03, to help in discussion by WG2 and UTC experts. The glyph for this character in Unicode 9.0 (June 2016) was derived from the font used to typeset the 2008 edition of the Tangut-Chinese Dictionary (*Xià-Hàn zìdiǎn* 夏漢字典) of Prof. Lǐ Fànwén 李範文. In Unicode 14.0 (September 2021) the glyph was modified to follow the form shown in the draft Table of Xixia Characters (西夏字符及属性标注表（草稿）) issued by the Zhonghua Ziku project research group for collecting and arranging Tangut Characters (中华字库工程西夏文字的搜集整理课题组) in August 2016.

U+17F03 in all versions of the Unicode Standard

9.0 through 13.0	14.0 through 17.0β
刭	刭

Comment CN1 for the CD.2 (consultation) to ISO/IEC 10646 7th edition (see WG2 N5316) requests that the glyph for U+17F03 be changed to 刭, which is the same shape as the component form of this character, and used in Hán Xiǎománg's 2021 Tangut dictionary, but is different to the form of the character used in the Unicode 9.0 through 13.0 code charts.

For reference, the glyph forms of U+17F03 given in various modern sources are listed in Table 1, and examples from a wide range of primary sources from the Western Xia (1038–1227), Yuan (1271–1368) and Ming (1368–1644) periods are listed in Table 2.

Extracts from previous WG2 documents discussing this character are given on page 6, and an extract from the 2016 draft Table of Xixia Characters is given on page 7.

Table 1: Glyph forms of U+17F03 in various modern sources

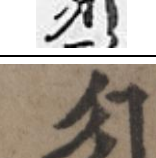| Source	Year	Glyph
韩小忙：《西夏文词典（世俗文献部分》 Hán Xiǎománg: *Xīxiàwén cídiǎn (shìsú wénxiàn bùfēn)* [Tangut Word Dictionary (Secular Literature Part)]	2021	
《西夏字符及属性标注表（草稿）》 *Xīxià zìfú jí shǔxìng biāozhù biǎo (cǎogǎo)* [Table of Xixia characters and annotated properties (draft)]	2016	
李範文、賈常業：《夏漢字典》 Lǐ Fànwén and Jiǎ Chángyè: *Xià-Hàn zìdiǎn* [Tangut-Chinese Dictionary]	2008	
Е. И. Кычанов, С. Аракава: *Словарь тангутского (Си Ся) языка. Тангутско-русско-англо-китайский словарь* E. I. Kychanov and Arakawa Shintarō: *Tangut dictionary. Tangut-Russian-English-Chinese dictionary*	2006	
韓小忙：《西夏文正字研究》 Hán Xiǎománg: *Xīxiàwén zhèngzì yánjiū* [A Study on Xixia Orthography]	2004	
李範文：《夏漢字典》 Lǐ Fànwén: *Xià-Hàn zìdiǎn* [Tangut-Chinese Dictionary]	1997	
李范文：《同音研究》 Lǐ Fànwén: *Tóngyīn yánjiū* [Study of the Homophones]	1986	
史金波、白滨、黄振华：《文海研究》 Shǐ Jīnbō, Bái Bīn, Huáng Zhènhuà: *Wénhǎi yánjiū* [Study of the Sea of Writing]	1983	
М. В. Софронов: *Грамматика тангутского языка* M. V. Sofronov: *Grammatika tangutskogo yazyka* [Grammar of the Tangut Language]	1968	
西田龍雄：『西夏文字小字典』 Nishida Tatsuo: *Seikamoji shōjiten* [Little Dictionary of Tangut Characters]	1966	

Source	Year	Glyph
Н. А. Невский: *Тангутская филология: Исследования и словарь* N. A. Nevsky: *Tangutskaya filologiya: Issledovaniya i slovar'* [Tangut philology: researches and dictionary]	1960	
Gerard Clauson: *Skeleton Tangut (Hsi Hsia) Dictionary*	1938–1939	

The form of U+17F03 requested by China matches the glyph used in Chinese publications during the 1980s and 1990s, and in the 2004 PhD dissertation and 2021 Tangut dictionary of Hán Xiǎománg. The Unicode 9.0 glyph matches the forms used in the Tangut dictionaries of Kychanov (2006) and Lǐ Fànwén (2008). Early authors such as Clauson (1938–1939) and Nevsky (1960) use a form that corresponds to that shown in the *Homophones* A text (see Table 2 no. 2).

Table 2: Examples of U+17F03 in primary sources

No.	Source	Glyph
1.	*Sea of Writing* (㗻㗱) [X] 1:70.271* (NB no head entry for this character survives)	
2.	*Homophones* (㗻㗱) A version [X] 04A27	
3.	*Homophones* (㗻㗱) B2 version [X] 39A11	
4.	*Homonyms* (㗻㗱㗱㗱㗱㗱㗱㗱) [X] 14A6 (character is only partially preserved)	
5.	*Pearl in Palm* (㗻㗱㗱㗱㗱㗱㗱㗱㗱) A edition [X] 17.6C	
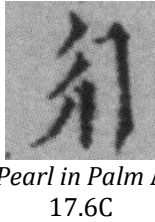
No.	Source	Glyph
6.	*Synonyms* (繆毿) [MS] 02A4	
7.	*Forest of Categories* (龘骸) [X] 3:14B5	
8.	*Sea of Meanings* (夠虤繆箾) [X] 10:01B8	
9.	*Grains of Gold* (蘳豥茓巍孩) [MS] 03B3	
10.	*Proverbs* (姕鬏) [X] 02B1	
11.	*Odes* (烲) [X] 1:03A6	
12.	*Tiansheng Law Codes* (皷散继藗憵陔縗骸) [X] 42B1	
13.	*Parental Love and Filial Piety* (厃軒蒲) [MS] §22	
14.	*Shining Speech of Three Generations* (散觚茘儭瓬魝孩) [M] 18A3	
15.	*Repentance Ritual of Great Compassion* (厃羴菠羴觔僟禭) [X] ch. 8	

No.	Source	Glyph
16.	*Thousand Names of the Buddhas of the Present Kalpa* (縊緀夊蘊孭絤縐溕菠) [X] §7	
17.	*Golden Light Sutra* (蕣贁傷繿肺席溕菠) [X] ch. 5	
18.	*Diamond Sutra* (蕣羲讀姦縐麀螽頽溕菠) [X] §26	
19.	*Lotus Sutra* (蕣羲讀姦縐麀螽頽溕菠) {Princeton University} [X] vol. 4 ch. 8	
20.	*Lotus Sutra* (蕣羲讀姦縐麀螽頽溕菠) {Jagiellonian University} [MS] vol. 5 ch. 14	
21.	*Heart Sutra* (絑縐勽絑菼讀姦縐麀螽頽絣溕菠) British Library Or.12380/2474 [MS] fragment	
22.	*Flower Garland Sutra* (鏺絾縗絑縗炏溕菠) [M] vol. 77	
23.	*Liangzhou Stele Inscription* (祥尾隤鏺貁绊炏縊織緈羢脪雒荄) {1094} [E] col. 9	
24.	*Eulogy of the Monument* (Juyongguan arch east wall) {1345} [E] col. 5	
25.	Bǎodìng dhāraṇī pillar A {1502} [E] side 7	

- Sources marked [X] are xylographic printed editions; sources marked [M] are movable type printed editions; sources marked [MS] are manuscript texts; and sources marked [E] are epigraphic inscriptions on stone.

Nineteen of the twenty-five examples of U+17F03 shown in Table 2 match the current code chart glyph; only two examples (nos. 1 and 12) match the Unicode 9.0 code chart glyph; and four examples (nos. 14, 15, 21, and 24) match the glyph form requested by China.

Extract from WG2 N5305

084. U+17F03

No.	Unicode 17.0α			XXWCD		
	Code Point	**Glyph**	**R/S**	**Vol./Page**	**Glyph**	**R/S**
084	U+17F03	刿	262.5	5/342	刿	LR141.5

Primary evidence:

not preserved *Homophones* A 38A75	*Homophones* A 04A27*	*Homophones* B2 39A11	*Pearl in Palm* A 17.6C	*Leilin* 3:14B5	*Shengli Yihai* 10:01B8	

Discussion and Notes: The glyph form of U+17F03 (no. 084) shown in *XXWCD* does occur in some sources, but the Unicode glyph form is much more common. See WG2 N5134 = L2/20-166 section 5.

Extract from WG2 N5134

Table 5. Summary of Glyph Corrections for 32 Tangut Ideographs

Code Point	13.0	New	Sample	Radical/Strokes		Notes
17F03	刿	刿		262.5	262.5	See "Table of Xixia Characters" p. 224

<div align="center">

《西夏字符及属性标注表（草稿）》

中华字库工程西夏文字的搜集整理课题组

二〇一六年八月

</div>

- The first column after the index number gives the 'orthodox character form' (正字), and the second column gives the 'variant character form' (异体字) or 'corrupt character form' (讹体字).

- The note against entry 4469 says "When used as a character component it is 刻, variant form of 刻". This difference is clearly illustrated in Fig. 1 below, showing the dedication pages of an edition of the *Lotus Sutra*, where U+17F03 as an ideograph (highlighted in red) is written as 刻, whereas as a character component (right part of U+182E8 㩲, highlighted in blue) it is written as 刻.