

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 kFanqie to Unihan Database
Source: Eiso Chan (陈永聪, Culture and Art Publishing House)
Reference: [L2/22-135](#), [L2/22-127](#)
Status: Individual Contribution
Action: For consideration by UTC
Date: 2022-12-28

0. Introduction

Fanqie (反切) is a traditional method to indicate the pronunciations of a syllable for Chinese character (Hanzi, Kanji, Hanja, Chŭ Hán) by using two characters and ending with “反” or “切”. The first character means the initial (聲母), which is called as “反切上字”, and the second character means the final (韻母) and tone (聲調), which is called as “反切下字”. In some early usages, the first character and the second character must share the same “等” and “開合”, but this practice had not been continued in the later time.

The following two tables show the fanqie analyses for U+6771 東 (德紅切) and U+4EBA 人 (如鄰切) based on 《廣韻》.

Table 0.1 Fanqie analysis for U+6771

UCS	Char.	fanqie	聲	攝	韻	等	開合	調
U+6771	東	德	端	曾	登	一	開	入
		紅	匣	通	東	一	開	平

Table 0.2 Fanqie analysis for U+4EBA

UCS	Char.	fanqie	聲	攝	韻	等	開合	調
U+4EBA	人	如	日	遇	魚	三	開	平
		鄰	來	臻	真 A	三	開	平

In the IRG encoding review works, the evidence for so many submitted characters shows the fanqie, or the submitted characters are used in a fanqie. The possible kFanqie property will be useful for the encoding works in future to get one form of Chinese readings to help the unification works, or judge if the unencoded character is the variant of encoded one when the character is only used in a fanqie based on the submitted evidence.

1. kFanqie

The proposed property values for **kFanqie** are collected from different versions of 《廣韻》 basically, which is the most important use of fanqie in the history, and I add some values for CJK Ext. H characters based on the IRG ORT. The current suggested values in the appendix are only collected and reviewed by me. It is better to be reviewed by other encoding experts later.

We could add more values based on other sources in future if possible. The fanqie used for Chinese dialects and other languages are all OK to include in kFanqie property. If the characters used in the suggested values are unified with the characters used in the existing values, or they are the undoubted variants anytime and anywhere, the suggested values will not be accepted.

All the fanqie in the real publishing sources are ended with “反” or “切”, and “反” and “切” could be interchangeable, so there is no need to include these two characters in the kFanqie property, that means all the values include and only include two CJK Unified Ideographs.

As Fig. 1.1 shows, the fanqie of 東 is 德紅切; as Fig. 1.2 shows, the fanqie of 東 is 德紅反. We should treat them as the same fanqie, that is the reason why I propose to include and only include two Unified Ideographs in the property values without ending with “反” or “切”. Therefore, the kFanqie property value of 東 (U+6771) should be 德紅. In fact, some characters have multiple readings, and hence multiple fanqie, so the values should be separated by space.

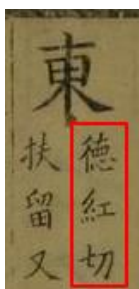


Fig. 1.1 陳彭年, 邱雍 等: 《廣韻》, 符山堂藏板, 上平聲卷第一



Fig. 1.2 吕大臨: 《考古圖釋文》, 四庫全書本, 上平聲

The following table provides the fields for this property as reflected in UAX #38.

Field	Text
Property	kFanqie
Status	Provisional
Category	Readings
Introduced	TBD
Delimiter	space
Syntax	<code>[\x{3400}-\x{4DBF}\x{4E00}-\x{9FFF}\x{FA0E}\x{FA0F}\x{FA11}\x{FA13}\x{FA14}\x{FA1F}\x{FA21}\x{FA23}\x{FA24}\x{FA27}-\x{FA29}\x{20000}-\x{2FFFD}\x{30000}-\x{3FFFD}]{2}</code>
Description	The fanqie reading, which is expressed as two CJK Unified Ideographs. The first one means “initial” (聲母), and the second one means “final” (韻母) and the tone (聲調). Complete fanqie readings always end with U+53CD 反 or U+5207 切 as a third character, which are interchangeable and therefore not included as part of the property value.

2. Coverage

A total of 20,261 CJK Unified Ideographs are assigned this proposed property. The following table indicates the number of CJK Unified Ideographs that are assigned this proposed property based on Unicode Version 15.0. The characters in Ext. C, D, E, F, G and the compatibility characters have not been assigned this proposed property currently. Any expert could request to add the property values for them in future.

	URO	Ext. A	Ext. B	Ext. H
Characters	13,387	2,697	4,154	23
Percent	63.77%	40.91%	9.72%	0.55%

3. Data file

The kFanqie.txt data file, which is a PDF attachment, provides everything that is necessary for adding this property to the Unihan Database, and for adding a full description of the property to UAX #38, Unicode Han Database (Unihan). The examples are shown as below.

U+4EBA kFanqie 如鄰
U+6771 kFanqie 德紅

In the data file, the number of the value(s) for one character is 1 through 7. The following table shows the situations.

	1	2	3	4	5	6	7
Characters	15,448	3,845	797	145	23	2	1
Percent	76.25%	18.98%	3.93%	0.72%	0.11%	0.01%	0.00%

In the current values, 483 characters are used as the first characters, 1,184 characters are used as the second characters, and totally 1,471 characters are used as the value characters.

For U+24865 (獮), I listed two values in the preliminary proposal, but one has been removed in this document. IRG had decided to encode [WS2021-02424:GKJ-00588](#) (𪛗 𪛘) separately during IRG #59, so the removed value should be moved to WS2021-02424 in future.

4. Bibliography

- 陳彭年, 邱雍 等: 《鉅宋廣韻》, 宋乾道五年刻本
陳彭年, 邱雍 等: 《廣韻》, 符山堂藏板
陳彭年, 邱雍 等: 《廣韻》, 龍谷大學藏至正南山書院刊本
陳彭年, 邱雍 等: 《廣韻》, 覆元泰定本
陳彭年, 邱雍 等: 《廣韻》, 澤存堂本
陳彭年, 邱雍 等: 《廣韻》, 古逸叢書覆宋本重修
陳彭年, 邱雍 等: 《廣韻》, 四部叢刊巾箱本
呂大臨: 《考古圖釋文》, 四庫全書本
IRG ORT

5. Acknowledgement

Prof. Hou Xingquan (侯興泉) provided some helps on this document.
Dr. Ken Lunde (小林劍), Mr. Lee Collins (康立論) and Mr. Stephan Stiller (尹賢雋) provided some comments.

(End of Document)