

Title: Proposal to add one ideograph to UAX #45 and to change the status of an existing UAX #45 ideograph

Source: Jerry Rossignuolo

L2/21-149

Status: Individual Contribution

Action: For consideration by the UTC

Date: 2021-08-06

This document is a proposal to 1) add one new ideograph to UAX #45 for submission to a future IRG working set; and 2) change the status of an existing UAX #45 ideograph so that it can be submitted to a future IRG working set. The primary basis of this proposal is that the ideographs are treated as core Indexing Components in 新华字典 (*XinHua Dictionary*) and in a number of official Chinese government publications.

The latest *XinHua Dictionary* utilizes 201 Principle Indexing Components and 79 Associated Indexing Components along with eight variant ideographs that are associated with five of the Principle Indexing Components. Of these 288 total ideographs, 285 of them are already encoded in the latest version of the Unicode Standard (Version 13.0). Among the three that are not encoded in the Unicode Standard, one may result from the future disunification of U+5F50 𠂔^[1], another one is already included in UAX #45 as UTC-00792 with status W, and the remaining one is the one that I am proposing to be added to UAX #45. Encoding these three ideographs would mean that the entirety of the ideographs used as Indexing Components in the *XinHua Dictionary* would be encoded in the Unicode Standard. In 2016, Guinness World Records officially confirmed that the dictionary, published by The Commercial Press, is the "most popular dictionary" and the "best-selling book (regularly updated)." It is considered a symbol of Chinese culture.^[2]

Furthermore, the two ideographs in this proposal are identified as fundamental Indexing Components of the current simplified Chinese writing system in a number of official Chinese government publications written explicitly for the purpose of standardizing Chinese Indexing Components.

The ideograph 𠂔, whose UAX #45 U-Source ID is **UTC-00792** and which is a variant of 车, is written using different strokes than 车. The ideograph 𠂔 is written with 提 (Upward Horizontal Tí) as its fourth stroke, while 车 is written with 横 (Horizontal Héng) as its fourth stroke. **The proposal is to change the UAX #45 status of UTC-00792 from W (not suitable for encoding as a CJK Unified Ideograph) to N (planned to be submitted for a future extension) so that it can be submitted to a future IRG working set.**

The ideograph 𠂔, which is a variant of 𠂔, is written differently than 𠂔, and differently than 𠂔 or 𠂔, which are two encoded variants of 𠂔. **The proposal is to add this ideograph to UAX #45 with status N.**

The table below lists the two ideographs and their metadata:

Ideograph	Variant	kRSUnicode	IDS	kTotalStrokes	First Residual Stroke	UAX #45 U-Source ID
𠂔	U+8F66 𠂔	159'.0	𠂔	4	0	UTC-00792
𠂔	U+807F 𠂔	129.0	𠂔一	6	0	N/A

I'd like to thank Ken Lunde for his thoroughly professional and positive support in helping me with this proposal, along with helping me with other questions about ideographs and the Unicode standard.

Evidence

The ideograph 车 is the core Associated Indexing Component number [68] of 新华字典 (XinHua Dictionary). It is distinct from the Principle Indexing Component number 68 车. See Figure 1.

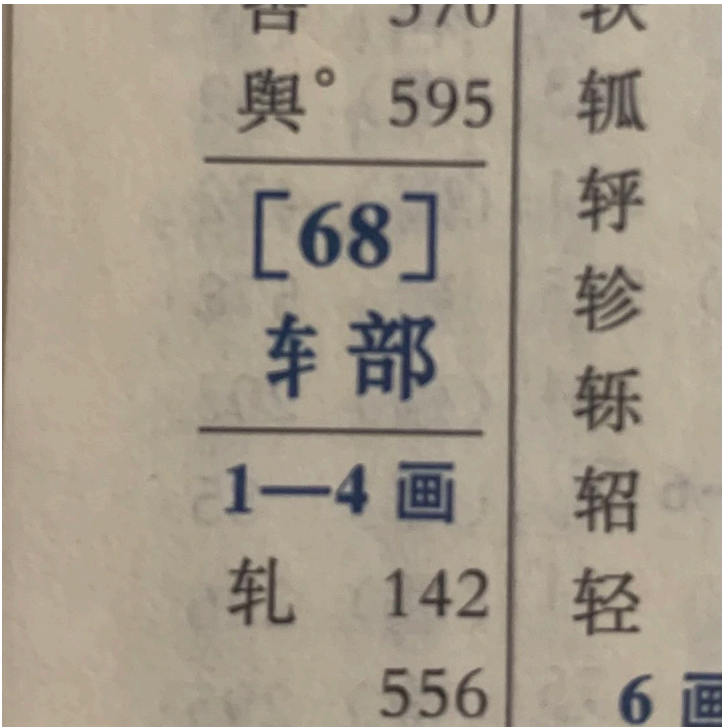


Figure 1 (Can be found in the XinHua Dictionary Version 12 copyright 2020)

The ideograph 车 can be found in the 新华字典 mobile app. See Figure 2.



Figure 2 (XinHua Dictionary mobile app Version 2.5.2)

The ideograph 𨍎 can be found on page 50 of the GF 0012-2009 standard, which is entitled GB13000.1 字符集汉字部首归部规范 (*Specifications for Identifying Indexing Components of GB 13000.1 Chinese Characters Set*). <http://ling.whu.edu.cn/law/002/2016-04-20/1307.html>. See Figure 3.



Figure 3

The ideograph 𨍎 can be found on page 3 of 汉字部首表 (*The Table of Indexing Chinese Character Component*) document GF0011-2009 published by the 中华人民共和国教育部国家语言文字工作委员会 (*National Language Working Committee of the Ministry of Education of the People's Republic of China*). See Figures 4 and 5.

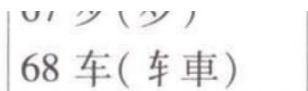


Figure 4

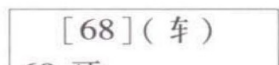


Figure 5

The ideograph 聿 is the core Associated Indexing Component number [145] of 新华字典 (XinHua Dictionary). It is distinct and written differently from the Principle Indexing Component number 145 聿. See Figure 6.

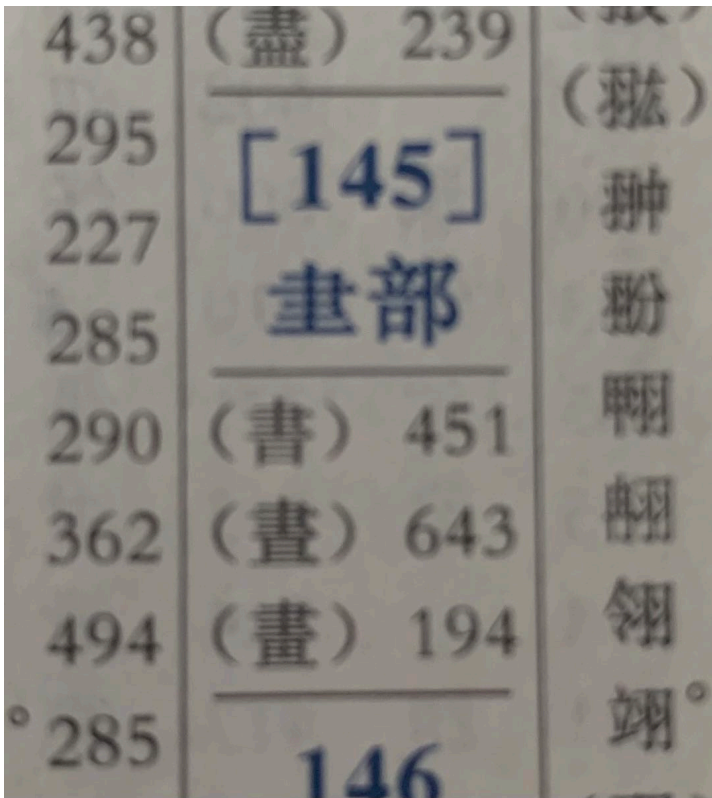


Figure 6 (Can be found in the XinHua Dictionary Version 12 copyright 2020)

The ideograph 聿 can be found in the 新华字典 mobile app. See Figure 7.



Figure 7 (XinHua Dictionary mobile app Version 2.5.2)

The ideograph 聿 can be found on page 8 of the publication GB13000.1 字符集汉字部首归部规范 (Chinese Character Component Standard of GB 130000.1 Character Set for Information Processing). <http://www.moe.gov.cn/ewebeditor/uploadfile/2015/01/12/20150112165337190.pdf>. See Figure 8.

196		(聿)	聿
197	114	聿	律 肇 盡
198		(聿)	書 畫
199		(聿)	畫

Figure 8

The ideograph 聿 can be found on page 27 of *The Writing Order of Modern Chinese Character Components* published in JMCLE Vol.2 No.1 May, 2013. <http://xuebao.eblcu.com/xuebao/essay/lixiaotong.pdf>. See Figure 9.

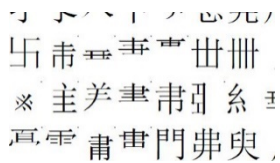


Figure 9

[1] <https://www.unicode.org/L2/L2020/20106r-unihan-adhoc-utc163.pdf> on page 8.

[2] https://en.wikipedia.org/wiki/Xinhua_Dictionary.