

Universal Multiple-Octet Coded Character Set  
UCS

ISO/IEC JTC1/SC2/WG2/IRG N2340

Date: 2018-10-15

Source:	Chen Zhuang, Wei Lin Mei, TSENG Shih-Shyeng, and Eiso Chan
Title:	Request to discuss WHEN and WHY to use IVS/IVD
Meeting:	IRG #51, Hanoi, Vietnam
Status:	Member's Submission and Ideographic Experts
Actions required:	To be considered by IRG
Distribution:	IRG
Medium:	Electronic
Page:	4
Appendix:	None

During the review works of WS2015 and WS2017, some reviewers pointed out some of the China-submitted characters and TCA-submitted characters should be added to IVS/IVD. When we review and respond the current version of WS2017, we deeply feel the new unification rule or the IVS/IVD rule is not easy to understand for the submitters. After a long and serious discussion among the authors of this document, we can't reach a consensus. Therefore, we list our questions as below and request to discuss them among all the IRG experts widely before discussing the review and response documents of WS2017 v2.0 in IRG #51. We will follow the consensus of this issue after the necessary discussion.

### **1. Guideline for "When and Why to use IVS/IVD"**

**We request IRG to give clear rules as a guideline of using "IVS/IVD", rather than a personal subjective recognition.**

During the review of WS-2017, "IVS/IVD" was clearly overused as reviewing opinions. **We felt it necessary to discuss when and why it was recommended an ideograph need be treated by using IVS/IVD.**

Based on IRG PnP v11 (IRGN2310) and **in the case without compatibility possibilities**, the followings are our understanding of the use of IVS/IVD:

- (1) If a submitted new ideograph should be unified to another existing ideograph, the submitter could only "register the IVS of that ideograph to IVD".

- (2) “Treating a submitted ideograph by using IVS/IVD” is a final decision made by the submitter rather than the position of any reviewer.
- (3) For the unification review, a reviewer should provide the opinion of either “not unified” or “unified to which ideograph” with necessary information, rather than any else opinion such as something relative to “IVS/IVD”.

We would like to strongly suggest ignoring any review opinion relative to “IVS/IVD” until IRG have given the guideline of using “IVS/IVD”. The unclear or inexhaustive comments based on the personal subjective recognition are NOT helpful for our works in fact.

## 2. Variants does not equal to IVS/IVD

IVS/IVD is a solution of encoding a unified ideograph rather than used for every ideograph variant. Right now, more than 95% (even 99%) of submitted ideographs are variants, and they are rarely “Zhengti ideograph”. “Zhengti ideograph” is the representative of a group of ideographs (variants) with same meaning and same pronunciation. However, the submitters did not provide enough information of variants. Examples are as follows:

WS2017 Seq. No.	Glyph1	variant/simplified	Evidence	Serial No. of Dictionary of Chinese character variants	Glyph in Dictionary of Chinese character variants	Zhengti ideograph	Source document
02636	璿 V-F0361	璿 (U+74CF)		A02609-003	璿	璿	《宋元以來俗字譜·玉部》引〈太平樂府〉。
02732	畛 USAT07201	Similar and synonym: 畛 (U+24C4F)/ 畛 (U+755B)		B02726-004	畛	畛	《龍龕手鑑·田部》。 《中華字海·田部》。
02964	碯 GDM-00116	碯 (U+7911)		C08899-001	碯	碯	《佛教難字字典》。
03052	礧 UK-10012			C08890-002	礧	礧	《重訂直音篇》。 The components of an ideograph could be moved and become a

							variant, such as “峰” vs. “峯”, “秋” vs. “焮, etc.
03140	裕 KC-07840			A03734-006	裕	裕	《偏類碑別字·衣部·裕字》引〈唐石州刺史劉穆墓誌〉。 《佛教難字字典·衣部》。

The Ministry of Education of Taiwan has spent more than 30 years collecting historical and modern literatures and has become a complete dictionary of variants. According to the information in this dictionary, TCA submits the complete Zhengti-variant relationship of each ideograph. It cannot be coded by “Zhengti ideograph” alone. It does not conform to the IVS/IVD principle. Second, in the case of absence of Zhengti-variant relationship, most of related ideographs cannot be linked only by their shapes.

### 3. About IRGPnP 1.3b

PnP V.11

#### 1.3 Discouragement of new disunification request

There is a possibility of “pure true disunification” request. This is almost like the new source code separation request. This kind of request shall not be accepted disregarding the reasoning behind. Key difference between “TO BE DISUNIFIED” and “SHALL NOT BE DISUNIFIED is as follows.

a. If character pair is non-cognate (meanings are different), that pair of characters is TO BE DISUNIFIED.

**b. If a character pair is cognate (means the same but different shape), that pair of characters SHALL NOT BE DISUNIFIED.**

- (1) For clause 1.3b, does “cognate” mean “variant”? And does the clause means that two variants with different shapes also should be unified even if they are not unifiable according to current unification rules?
- (2) If the answer for the question above is “yes”, most of the ideographs submitted by each member are variants of some already encoded ideographs, but they are not marked with Zhengti-variant relationships. Even those ideographs without Zhengti-variant relationships, they are not surely Zhengti ideographs. Does the

work of IRG continue? The ideographs submitted in future must prove that it is a Zhengti ideograph for submission?

Note: The MOE's Dictionary of Chinese Character Variants contains 106,330 ideographs (29,921 Zhengti ideographs with their variants 74,407 ideographs) and 2,002 ideographs pending for further analysis. All of the 29,921 Zhengti ideographs had been encoded in UCS. If one can find any (neither personal name nor location name) Zhengti ideograph outside the MOE's Dictionary of Chinese Character Variants, it will be a great contribution.