

ISO-IEC JTC1/SC2/WG2/IRG

**Universal Multiple-Octet Coded Character Set**

**(UCS)**

JTC1/SC2/WG2/IRG **N 207** (R)

JTC1/SC2/WG2 **N 1236**

Date : 1995-06-26

**TITLE :** Status of Ideograph Vertical Extension

**Source :** Zhang Zhoucai, IRG Rapporteur

***Requested***

**Actions :** for WG2's Consideration at its 28th Meeting in Helsinki

Refer to : **WG2 27th Meeting Resolution M27.17 and M27.18**

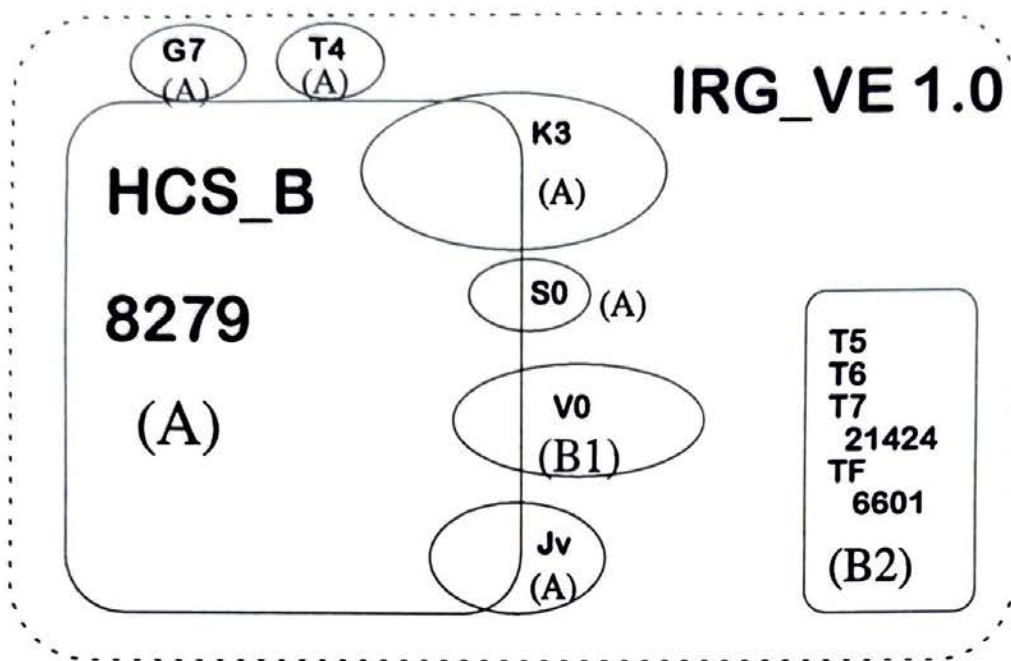
**Vertical Extension Summary**

**1. Ideographs**

According to IRG Project Editor's report, the data for vertical extension from China(G), Japan(J), Korea(K), Vietnam(V), TCA(T) and Singapore(S) have been collected and merged preliminarily.

HCS\_B ( **8279** ideographs ) is used as the basis of unification of C, T, J, K, V and S, the newly formed subsets with HCS\_B are shown in the next page.

Now, this collection is renamed as IRG\_VE version 1.0 Draft, which contains  $8279 + 4203 = \mathbf{12482}$  ideographs of Category A and B1, as well as 28025 ideographs of Category B2.



**IRG VE Version 1.0 (Draft)**  
**HCS\_B part**

|                                                 | G3                                                                                                            | G5                                                                             | TE_Unique                                                                       |
|-------------------------------------------------|---------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| <b>TE</b>                                       | <b>0</b><br>Note: Actually, this subset has been included in ISO 10646.1. Here "0" implies none left for V.E. | <b>682</b><br>including<br>Jv 167<br>V0 46<br>K3 129<br>S0 10<br>( Unifiable ) | <b>1687</b><br>including<br>Jv 117<br>V0 25<br>K3 167<br>S0 14<br>( Unifiable ) |
| <b>Gn_Unique</b><br>( No intersection with TE ) | <b>2400</b><br>including<br>Jv 53<br>V0 35<br>K3 453<br>S0 6<br>( Unifiable )                                 | <b>3510</b><br>including<br>Jv 64<br>V0 48<br>K3 461<br>S0 9<br>( Unifiable )  |                                                                                 |

In addition to the above subsets in HCS\_B, there are several unique subsets ( unique part ):

G7 (see Remark) **185** ( A category ) with font and Kangxi data  
T4 **750** (A category) with font and Kangxi data  
K3 2149-1210= **939** (A category) with font and Kangxi data  
S0 328-39= **289** (A category) with font, lacking Kangxi data  
V0 1775-129=**1646** (B1) with font in special format , and Kangxi data.  
Jv 678-284= **394** (A category) with font and complete Kangxi data

**SubTotal 4203 ( A + B1 / Unique )**

T5 +T6+T7 21424 (B2)  
TF 6601 (B2)

**SubTotal 28025 ( B2 )** - Not candidates for inclusion in BMP

**Remark:** The above 185 Hanzi , temporarily coded in G7, are from

(1) The Summary Table of Simplified Hanzi 《简化字总表》

**60** simplified Hanzi and their unsimplified counterparts.

(2) **38** additional radicals found in 5 Asian dictionaries.

( 《康熙字典》 , 《汉语大字典》 , 《大夏和辞典》 ,  
《大字典》 and 《辞海》 )

(3) **87** contemporary Hanzi , most of which are simplified counterparts of some traditional ideographs coded in ISO 10646.1.

## **2. Prioritization of the Ideographs in Vertical Extension**

Since there was no IRG meeting between WG2 27th meeting and 28th meeting, IRG members have no chance to meet together to discuss on prioritization of ideographs in Vertical Extension. The current priorities of the ideographs in vertical extension shown in 1.1 is categorized as A, B1 and B2, after freezing source codes for vertical extension. However, the unification of these ideographs will be further refined , the final repertoire to be coded in BMP is expected to be submitted to 29th WG2 meeting in Tokyo.

**Rapporteur's Announcement:**

According to IRG/N189 Conclusion V2, and the instruction in WG2/N1204 Resolution M27.18, ideograph source repertoire for consideration of IRG 's extension proposal for the second edition of ISO/IEC 10646 **has been frozen** after due consideration of Singapore's input.

**3 Components of Ideographs**

The idea of components of ideograph to be coded in BMP , was accepted by WG2/IRG, though it was recommended to be discussed separately from vertical extension of ideographs.

The number of components of ideograph is estimated to be less than **500**. More than 40000 ideographs can be represented using these components .

**Rapporteur's Note:**

*China's proposal on Structure Symbols for combining method of ideograph representation is still under discussion in IRG. The result of the discussion will be reported to WG2 after the 5th IRG meeting.*

**Contact:** Zhang\_Zhoucai 张周才 MEI/CCID

Rapporteur of ISO-IEC JTC1/SC2/WG2/IRG

54 Xisanhuan Bei Rd. Beijing 100044

TEL: (8610)841.2606 or 847.1418

FAX: (8610)841.1363 E-mail: **ZhangJoe@sun.ihep.ac.cn**