

Date : 1995-08-18

ISO-IEC JTC1/SC2/WG2/IRG
Ideographic Rapporteur Group
 Secretariat: China

TITLE: Status of Vertical Extension
SOURCE: Chief Editor
ACTION: Discussion at IRG 5# Meeting
Refer to: IRG N207R

1. Source Repertoire for the Vertical Extension

SET	G7	G3	G5	TE	T4	T5	T6	T7	TF	JA	K3	V0	S0
Category	A	A	A	A	A	B2	B2	B2	B2	A	A	B1	A
Amount	42+	2400	4192	2368	749	1174	1526	520	6601	678	2149	1775	328
	75				-1				-1				-16

Note1: Amongst G7, G3 and G5, there should not be any intersection.

Note2: Amongst T4, T5, T6, T7 and TF, there should not be any intersection.

2. Problems with Some Source Repertoire

2.1 Duplication with ISO 10646/CJK Unified Ideographs

S0: 16 Hanzi exist in CJK

T6: T6-264E with U+5438; T6-2131 with U+4E38

JA: JA-2373 with U+8012

K3: K3-224F with U+867D

2.2 Duplication within the same source

T4 with TF:

4-2226 with F-216D

4-2A2D with F-2C27
 4-3625 with F-392B
 E-6322 with 4-302C (HCS_B B941)
 E-2354 with F-2249 (HCS_B A4FB)

- 2.3 Incorrect Kangxi version or Un_normal Kangxi data
 V0: Not the 7th Version Published in Beijing.
 K3: The virtual Kangxi code appears as "2" or "3".
- 2.4 Lacking Kangxi data
 S0: Filled by CCID, not reviewed yet.
- 2.5 Data records appear twice
 T4-6E4D and TF-5E5B
- 2.6 Late submission of electronical data and font
 V0.24 in correct MRF was received on 8 August
 Source code in correct MRF was received on 11, August

3. Merging Statistics for Vertical Extension Draft 2.0

Unified VE Core: Amongst G,T,J,K,S and V,
 at least two of them have intersection in the core.

	G7	G3	G5	T with no G
TE	5	0	682	228+ 1453
T4	0	6	9	73+ 660
T5	1	6	11	33+ 1123
T6	1	4	4	27+ 1490
T7	0	0	2	9+ 509
TF	2	6	12	54+ 6526
G with no T	9+	470+	472+	
	99	1908	3000	

Not unifiable with
 G,J,K,V or S

Not unifiable with T,J,K,V or S

Intersections of
 J, K, V and S
 (Except J/K/V/S-Unique)

In addition to the above Unified VE Core including the intersections of J,K,S and V, as well as G/T-Unique subsets, there are 4 more unique subsets as follows:

JA-Unique: 296
 K3-Unique: 946
 V0-Unique: 1773 (Not ready to unify yet)
 S0-Unique: 232

Appendix: Subsets formed by G,T,J,K,V and S source repertoire

- (1) The number of subsets formed by G and T only:

$$\overset{1}{C}3 + \overset{1}{C}6 + \overset{1}{C}3 \overset{1}{C}6 = 3+6+3 \times 6=27$$

- (2) The number of the subsets formed by J, K, V and S :

$$\overset{1}{C}4 + \overset{2}{C}4 + \overset{3}{C}4 + \overset{4}{C}4 = 4+6+4+1=15$$

- (3) The number of all subsets formed by G,T,J,K,V and S together:

$$27 + 15 + 27 \times 15 = 447$$

- (4) The vector "GTJKVS" can be used to describe the 447 subsets by assigning GTJKVS with different values. Here,

G= 7, 3, 5, *

T= E, 4, 5, 6, 7, F, *

J= A, *

K= 3, *

V= 0, *

S= 0, *

The "*" indicates there is none from that source repertoire in a specific subset.

For example, GTJKVS=3FA300 shows the subset is the intersection of G3, TF, JA, K3, V0 and S0; or, for a specific ideograph, it implies the ideograph is from G3, TF, JA, K3, V0 and S0. If a subset is T4-unique, it can be denoted as GTJKVS=*4****.

For details, see Attachment to IRG N226.

Contact:

Zhang_Zhoucai 張軸材 /Rapporteur of IRG

FAX: (8610)841.1363

E-mail: ZhangJoe@sun.ihep.ac.cn