

ISO/IEC JTC 1/SC2/WG2/IRG Ideographic Rapporteur Group

Source / Contribution Identifier: Japan

Meeting: IRG#19 @ Macao

Title: Comments on IRG N856

Keywords: (none)

Status: Member body contribution

Short Description: Japan Comments on IRG N856 (Not Encoded 89 Hanzi In KangXi Dictionary)

Proposed Conclusion / Requested Action: To be discussed by editors

Arguments / Text of Contribution: See below.

0. Introduction

The document IRG N856 lists 89 Kangxi *hanzi* and states they are “not encoded.” Review showed that the statement is partly correct and partly wrong. It also raised some other issues.

1. Some characters are already encoded in UCS

Some characters are found in existing UCS standard. They are shown below.

KXI	SCJK	UCS
0105.190	10701	00021940
0111.300	36895	00025741
0396.270	16432	00006132
0463.030	19407	00003A6E
0905.040	39569	00004270
1002.060	08043	0002132B
1226.030	54290	00028029
1381.200	12066	00021D11
1381.210	12064	00021D0F
1381.220	12065	00021D10
1396.140	62632	0002940A *
1551.300	10392	00021862
1569.600	55785	00028400
1585.110	00263	000200A7
1585.130	00296	000200C4
1586.460	01979	000203FD
1588.230	10724	0002194B *
1592.460	09328	0002166B
1595.800	13594	000220AF
1598.290	16015	000226AA
1599.790	19721	00022F37 *
1605.380	29984	00007280
1609.530	18506	00022C3F *
1612.270	07213	00021156 *
1613.410	43355	0002668D

*) Application of the unification rules to these characters is unassured. To be checked by Editors.

2. Some characters are in UCS and SuperCJK shows their shapes inappropriately

Two more characters are found in UCS. However, a reviewer also found that SuperCJK 14.0 shows their shapes inappropriately.

The first one is:

KXI	SCJK	UCS
0583.240	24705	00006BA9

Although the standard (ISO/IEC 10646-1:2000) shows shapes that are almost same as that in Kangxi, (for whatever reason,) the shape in SuperCJK 14.0 is different. Any shape in the standard does not look like SuperCJK. The shape (font) in SuperCJK should be corrected.

The second one is:

KXI	SCJK	UCS
0547.170	22983	000069F1

This case is more complicated. The standard shows several shapes for 000069F1; one of it just looks like Kangxi, and another just looks like that in SuperCJK 14.0. Unfortunately, Japan has a feeling that they should not be unified. We need a discussion here.

3. Not found in UCS

Some characters are not found in UCS. So the title of N856 is correct for these characters. However, there are more issues.

3.1 KXI is in SuperCJK with G_KX source indication

The following KXI are found in SuperCJK 14.0. For the corresponding code points, SuperCJK shows shapes that are not to be unified with those in Kangxi. However, those code points also have G_KX source indication. It must not happen.

Since these code points also have non-G sources, appropriate editors should review the corresponding font. If the shown fonts are correct, we suggest removing G_KX source indication from those code points.

KXI	SCJK	UCS
0129.220	07504	00021218
0349.260	03933	00020919
0539.240	22633	00023551
0560.150	23664	00003C08
0563.270	23906	00023839
0579.360	24581	00023A10
0580.150	24543	000239F7
0591.130	24947	00023B0E
0686.180	29280	00024454
0749.040	32087	00024B3E
0753.090	32270	00024BAD
0762.030	32548	00024C67
0852.250	37069	0002579B
0885.160	38593	00025B5C
0887.290	41875	000262CA
0928.110	43421	000266A9
1047.080	46178	00026CED
1062.250	47195	00026F5F
1098.280	49230	00027466

1454.160	65596	00029B3B
----------	-------	----------

3.2 KXI is in SuperCJK with G_KX single source indication

The following KXI are found in SuperCJK. The shape shown in SuperCJK 14.0 are not to be unified with those in Kangxi. However, SuperCJK also indicates that they are G_KX single source. That means either:

- (1) If the source information indicated in SuperCJK is true, the shape shown in SuperCJK (and in ISO/IEC 10646-2:2001) are wrong, and they should be corrected.
- (2) If the shapes shown in SuperCJK are correct, the source information must be incorrect. Although they are single-source, there must be other source... (What is it?)

We need a discussion here.

KXI	SCJK	UCS
0652.280	27230	00023F72
1185.340	52236	00027B66

3.3 KXI is in SuperCJK but are not marked as G_KX

The following KXI are found in SuperCJK. This is not a problem. The UCS code points indicate they are pre-Extension-B characters, and the KXI of those days was just a reference to a character in Kangxi (of same meaning and/or origin) that may or may not be unifiable.

KXI	SCJK	UCS
0239.130	08509	000058AF
0398.090	16485	0000614C
0928.010	40685	00007DB2
1061.190	47150	00008598
1249.170	55694	00008F5A
1376.170	61498	0000971A

3.4 KXI is not found in SuperCJK

The following KXI are not in SuperCJK. They look just left behind.

0079.090	1174.260
0081.140	1266.260
0086.020	1393.070
0124.070	1393.260
0162.240	1548.690
0335.010	1553.560
0369.240	1563.260
0401.310	1570.430
0422.030	1570.710
0472.270	1591.a40 *
0480.110	1602.220
0547.100	1606.960
0584.120	1608.990
0589.070	1614.430
0637.080	1626.040
0642.170	1630.360
1123.330	1631.110

*) I don't like the idea of writing "104th character on page 1591" as 1591.a40, although I find no better alternative.

(END OF DOCUMENT)