1. Background:

- IRG PnP v60 Draft4 (IRG N1952) has the following clause RE: evidence required for CJK characters to be encoded in UCS.

2.2.3. Required Evidence to be Submitted

a. Supporting Evidence: Evidence of the proposed glyph shape, its usage and context with pronunciations, meanings, etc. should be supplied to convince the IRG that it is actually used or non-cognate with other similar ideographs. Evidence for each character must be supplied as scanned images. The provision of evidence on character usage including those for personal names should not be exempted. A declaration for character use without accompanying evidence is generally not acceptable. Considering privacy issues, the IRG has suggested some compromised provisions. Details are given in Annex G Part 3.

Note: To support e-government related initiatives, IRG can at its discretion accept submissions of characters that are used in computer systems administered by government bodies for public service with wide access by government agencies and citizens. Factors considered for such acceptance are further elaborated in Annex G.4.

G.4. Consideration for acceptance for characters that cannot be provided in Printed forms

The consideration for acceptance for characters that cannot be provided with evidence in printed forms is not meant to relax the IRG requirement on the provision of evidence for modern use. Rather, it is meant to facilitate e-government initiatives for computerized processing of information. It is
under this presumption the IRG is willing to consider acceptance of characters already supported in computer systems that are maintained by a designated government body with wide use by government bodies and citizens for administrative public service.

The IRG recognizes that due to e-government issues, some of the characters included in such systems cannot supply supporting evidence for actual use according to 2.2.3a in this document, yet it is technically and administratively not practical to remove them from the computer systems. Thus IRG is willing to consider their acceptance without actual evidence for already implemented working system only. However, IRG requires the submitting member bodies to provide information on the quality assurance process for the maintenance of the respective character collections. Submitters must supply information on the accessibility of the respective character collection and the respective system, the stability and the traceability of the collection, and the kind of evidence/information needed for approval of character removal, modification and addition by the administrative body of the collection.

2. Analysis

2.1 Hanja chars NOT USED for personal names are included in the list for personal names

- ROK government has a list of tens of thousands of Hanja characters for personal names.
- However, nobody seems to know exactly what characters are “actually” used for personal names and what are “not”.

2
2.2 How to proceed to encode Hanja chars for personal names in UCS?

- If ROK IRG committee is to encode these Hanja characters in UCS, the committee might will probably proceed as follows:

  a) There is a list of tens of thousands of Hanja characters (List F: Full) for personal names maintained by ROK Government. Each character in List F "may or may not be actually" used for personal names.

  b) Out of List F, make a list of Hanja characters (List AU: Actually Used) that are "ACTUALLY" used for personal names.

  c) Out of List F, make a list of Hanja characters (List NU: Not Used) that are NOT actually used for personal names.

  d) Out of list AU, make a list of Hanja characters (List UH: UCS Hanja) that are already encoded in UCS and has glyphs in K column. Neither horizontal nor vertical extension is needed for these Hanja characters.

  e) Out of list AU, make a list of Hanja characters (List HE) for horizontal extension in UCS.

  f) Out of list AU, make a list of Hanja characters (List VE) for vertical extension in UCS.

2.3 The relationships among the above Lists

- The relationships among the above Lists can be summarized as follows:

  \[ F = AU + NU = UH + HE + VE + NU \]
  \[ AU = UH + HE + VE \]

- The relationships can be visualized in Venn Diagram as follows:
2.4 The main question: Do we really want to encode characters in List NU (NOT USED FOR PERSONAL NAMES) in UCS?

- ROK IRG committee considers that encoding Hanja characters in NU List (Not Used for personal names) is NOT desirable for ROK users or UCS users.

- ROK would like to hear opinions from MBs in IRG.

2.5 Situation in Japan?

- The situation in Japan might be quite similar to that in Rep. of Korea, except that Japan probably includes Kanji characters for place names in addition.

- ROK IRG committee expects that almost the same reasoning can be applied to the situation in Japan and therefore, a paper evidence for actual usage of each Kanja character must be supplied (as IRG has been doing until now) to encode that character in UCS.
2.6 Past Experience

- A few years ago, Macau submitted tens of Hanzi characters for UCS which, Maccau said, are used for personal names.

- When IRG asked for evidences and checked them, a few Hanzi characters were excluded due to the fact that there were no evidences for those chars.

- ROK considers that such a practice was really good to ensure the quality of CJK characters in UCS.

3. In the future

- If IRG really wants to encode Kanji characters in UCS per se IRG PnP 2.2.3.a and Annex G.4, ROK might take this issue to WG2 and get a final decision from WG2.

- If ROK should submit several thousands of Hanja characters for personal names, will they be accepted for inclusion in UCS “without evidences”? (Of course, thousands of Hanja characters will be divided into manageable pieces to be reviewed in each IRG project such as ExtG, ExtH, etc.).

* * *