

Subject: A Mapping table between WS SN (Serial Number) and UCS cp for WS2015
Date: 2021.03.08.
Author: KIM, Kyongsok (Republic of Korea)
Status: Individual Contribution

1) As you can see below, in the excel file IRGN2308WS2015v6.0Attributes2018-06.xlsx, currently there seems no direct mapping information between SN (Serial Number) in Working Set 2015 (which is now Ext. G) and UCS cp.

	A	C	E	F	H	J	L	S
1	Sequence Number	UTC Source	UK Source	T Source	K Source	SAT Source	G Source	IDS
2	00001		UK-02764					ㄱ
3	00002		UK-01671					ㄴ
4	00003	UTC-00969						ㄷ
5	00004						GHZR10110.1	ㅁ
6	00005					USAT05803		ㅂ
7	00006						G_Z3422507	ㅅ
8	00007		UK-02766					ㅇ
9	00008						GHZR10014.0	ㅈ

2) For example, to do IDS checking of WS2021 candidate chars, WS2015 SN can be used in ws2015 ids file. However, if we have UCS cp in addition, it will be a lot better for checking the UCS glyphs in ISO/IEC 10646 since we could use ExtG TTF font easily.

3) As another example, when IRG MBs submitted TTF font for WS2015, temporary or private UCS cp was used at the time of creating TTF fonts.

Later, we need to change temporary or private UCS cp's to the final UCS cp's in these TTF fonts. Once again, if we have the UCS cp in addition, the change will become a lot easier.

4) Based on the observations mentioned above, I created a mapping table among three data items: WS2015 SN, UCS cp, and source reference. Its document number is:

IRG N2465_2_ws2015_map_t_sn_ucs_cp_sr_v10.txt (# lines: 4997)

```

k2509_2_ws2015_map_t_sn_ucs_cp_sr_v10.txt - Windows 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)
00001 U+30000 UK-02764
00002 U+30001 UK-01671
00003 U+30002 UTC-00969
    
```

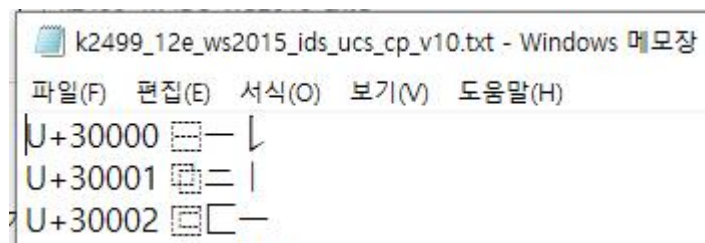
There is one line for each source reference and, therefore, if there is more than one source reference for some SN/UCS cp., there are two or more lines in the mapping table, as shown below. This explains why the number of lines in IRG N2465_2, 4997, is greater than the number of chars in ExtG (<- WS2015), 4939.



- I made some changes to create the mapping table correctly (hopefully), which required quite some time and effort to figure out problems and solve them. See Appenex 1 for details.

5) Using IRG N2465_2 in 4), I also created a ws2015 IDS table v1.0 using UCS cp (instead of WS2015 SN or source reference). Its document number is:

IRG N2465_3_ws2015_ids_ucs_cp_v10.txt (# lines: 4939)



6) Comments are welcome. Please let me (gimgs@pnu.kr) know if there is anything wrong in the IRG N2465_2_ws2015_map_t_sn_ucs_cp_sr_v10.txt or IRG N2465_3_ws2015_ids_ucs_cp_v10.txt

7) When CJKSrc.txt for ExtH becomes available, I may create an ExtH (WS2017) ids table using UCS cp (instead of WS2017 SN or source reference).

- I would like to suggest that the WS2017 attribute file be updated once WS2017 becomes ExtH. Specifically, two issues come to my mind. It took much time to fix these problems.

a) The number of SNs (i.e., lines) in the M-set of the WS2017 attribute file should be the same as the number of UCS cp's allocated to ExtH in CJKSrc.txt.

b) The values of the source references in the M-set of the WS2017 attribute file should be the same as the values of the source references in ExtH in CJKSrc.txt.

- Once the above problems are fixed, it is not hard to create a mapping table and an ExtH (WS2017) ids table.

* * *

See also

IRG N2465_2_ws2015_map_t_sn_ucs_cp_sr_v10.txt

IRG N2465_3_ws2015_ids_ucs_cp_v10.txt

k2509_1_N2465_1_map tbl_WS_sn__ucs_cp.hwp

Appenex 1. Lines deleted and added to create Mapping table and ids table for ExtG (WS2015)

Mapping table for ExtG (WS2015)

1) The following two lines are deleted from IRGN2308WS2015v6.0Attributes2018-06.xlsx and then, from this file, IRG N2465_2_ws2015_map_t_sn_ucs_cp_sr_v10.txt is created.

01436 UK-02790 -> horizontal extension at U+22C3A
02868 GZ-1861301 -> seems removed from ExtG at the last moment

2) The following three lines are added to IRG N2465_2_ws2015_map_t_sn_ucs_cp_sr_v10.txt

01638 U+30DD9 GZ-4511301 -> seems added to ExtG at the last moment
02231 U+307BF GZ-0722601 -> seems added to ExtG at the last moment
03617 U+30C5D UK-01969 -> seems added to ExtG at the last moment

Note. ExtG (WS2015) ids table

- Similar changes were made to create N2465_3_ws2015_ids_ucs_cp_v10.txt.

01436 UK-02790 removed
02868 GZ-1861301 removed

U+30DD9 暑負 added (GZ-4511301)
U+307BF 战 added (GZ-0722601)
U+30C5D ++项 added (UK-01969)

* * *