

IRG N2175

Title: Canonicalization for Variant-insensitive searching and collation

Source: Henry Chan

Type: Individual Contribution to IRG #47

Date: 2016/10/15

Introduction

Due to the nature of the Hanzi Script, there are often multiple variants of the same character in common use. Per the Source Separation Rule, various true duplicates and mis-disunifications, many similar shape glyphs for the same character have been encoded at different code-points in UCS.

For the average user, it may be puzzling to find that when searching for 水溫, a normal search engine will fail to return any entries for 水溫. In some other contexts, having an exact match may be more preferred, such as for household registration system.

Solution

The solution is to normalize the characters to a given variant/code-point. It is proposed that an SVS be registered for all identified similar-shape glyphs encoded at different code-points for a given character:

Example (1): 6B65 步 and 6B69 步

6B65 FE00 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-6B69 (步)

6B69 FE00 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-6B65 (步)

Example (2): U+5433 吳, U+5434 吳 and U+5449 吳

5433 FE00 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-5434 (吳)

5433 FE01 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-5449 (吳)

5434 FE00 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-5433 (吳)

5434 FE01 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-5449 (吳)

5449 FE00 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-5433 (吳)

5449 FE01 is registered to be equivalent to CJK UNIFIED IDEOGRAPH-5434 (吳)

For every n variants identified, $(n)(n - 1)$ SVSs will be registered.

During indexing and searching, the engine can convert the text to their normalized variants with SVS via a mapping table, then skip over SVS when doing comparison. Such a method is much more efficient than maintaining complex multiple-way conversions.

Sorting can also be carried out on the normalized form using SVS. In some locales, Radical/stroke ordering is preferred. Different variants may belong under the different radicals, making the list seem unordered. Different variants may also have different stroke counts. When using an SVS solution, the sorting engine can simply sort characters according to their normalized glyph shape, then sort by the specific variant.

As each region may have a different preferred glyph, it is recommended that for all variants identified, an exhaustive list of SRS for those variants are registered. Also, for some locales two different glyphs may have their own meanings, while for some locales two different glyphs are the same character, such as 么 幺 are two different characters in Chinese (Simplified), but regarded as variants in Chinese (Traditional). This will accommodate any possible changes and developments in preferred orthography.

These SVS mappings can be used to encode not only variants, but also true duplicates such as 烱 (U+3DB7) and 烱 (U+2420E), 栽 (U+363D) and 栽 (U+39B3), 豎(U+249BC) and 豎 (U+249E9) etc. For these exact duplicates, maybe only one mapping is necessary. The SVS can also serve to work around the problems caused by mis-disunified glyphs later identified.

There are in total 16 available SVS that can be registered. It is not expected that over 16 unifiable variants will be identified inside UCS for a given character. Once an SVS is registered to another variant, the same SVS for a given base character should not be reused in case of any error identified. The variant preferred for each locale may be stored in a central repository such as the CLDR.

A list of SVS to be registered for Source Code Separations has been kindly generated by Dr. Ken Lunde and attached as Appendix I. This list could be later updated with true duplicates identified inside UCS as well as other mis-disunified glyphs.

-- End of Document

Appendix I: StandardizedVariants.txt for Source Code Separations:

4E1F FE00; CJK UNIFIED IDEOGRAPH-4E22;
4E22 FE00; CJK UNIFIED IDEOGRAPH-4E1F;
4E48 FE00; CJK UNIFIED IDEOGRAPH-5E7A;
4E89 FE00; CJK UNIFIED IDEOGRAPH-722D;
4EDE FE00; CJK UNIFIED IDEOGRAPH-4EED;
4EED FE00; CJK UNIFIED IDEOGRAPH-4EDE;
4F75 FE00; CJK UNIFIED IDEOGRAPH-5002;
4FA3 FE00; CJK UNIFIED IDEOGRAPH-4FB6;
4FB6 FE00; CJK UNIFIED IDEOGRAPH-4FA3;
4FC1 FE00; CJK UNIFIED IDEOGRAPH-4FE3;
4FDE FE00; CJK UNIFIED IDEOGRAPH-516A;
4FE3 FE00; CJK UNIFIED IDEOGRAPH-4FC1;
4FF1 FE00; CJK UNIFIED IDEOGRAPH-5036;
5002 FE01; CJK UNIFIED IDEOGRAPH-4F75;
5024 FE00; CJK UNIFIED IDEOGRAPH-503C;
5036 FE00; CJK UNIFIED IDEOGRAPH-4FF1;
503C FE00; CJK UNIFIED IDEOGRAPH-5024;
5077 FE00; CJK UNIFIED IDEOGRAPH-5078;
5078 FE00; CJK UNIFIED IDEOGRAPH-5077;
507D FE00; CJK UNIFIED IDEOGRAPH-50DE;
50DE FE00; CJK UNIFIED IDEOGRAPH-507D;
514C FE00; CJK UNIFIED IDEOGRAPH-5151;
514E FE00; CJK UNIFIED IDEOGRAPH-5154;
5151 FE00; CJK UNIFIED IDEOGRAPH-514C;
5154 FE01; CJK UNIFIED IDEOGRAPH-514E;
5156 FE00; CJK UNIFIED IDEOGRAPH-5157;
5157 FE00; CJK UNIFIED IDEOGRAPH-5156;
516A FE00; CJK UNIFIED IDEOGRAPH-4FDE;
518A FE00; CJK UNIFIED IDEOGRAPH-518C;
518C FE00; CJK UNIFIED IDEOGRAPH-518A;
51C0 FE00; CJK UNIFIED IDEOGRAPH-51C8;
51C8 FE00; CJK UNIFIED IDEOGRAPH-51C0;
51E2 FE00; CJK UNIFIED IDEOGRAPH-51E3;
51E3 FE00; CJK UNIFIED IDEOGRAPH-51E2;
5203 FE01; CJK UNIFIED IDEOGRAPH-5204;
5204 FE00; CJK UNIFIED IDEOGRAPH-5203;
520A FE00; CJK UNIFIED IDEOGRAPH-520B;
520B FE00; CJK UNIFIED IDEOGRAPH-520A;
5220 FE00; CJK UNIFIED IDEOGRAPH-522A;
5225 FE00; CJK UNIFIED IDEOGRAPH-522B;
522A FE00; CJK UNIFIED IDEOGRAPH-5220;
522B FE00; CJK UNIFIED IDEOGRAPH-5225;

5238 FE00; CJK UNIFIED IDEOGRAPH-52B5;
5239 FE00; CJK UNIFIED IDEOGRAPH-524E;
524E FE00; CJK UNIFIED IDEOGRAPH-5239;
524F FE00; CJK UNIFIED IDEOGRAPH-5259;
5259 FE00; CJK UNIFIED IDEOGRAPH-524F;
525D FE00; CJK UNIFIED IDEOGRAPH-5265;
5265 FE00; CJK UNIFIED IDEOGRAPH-525D;
5292 FE00; CJK UNIFIED IDEOGRAPH-5294;
5294 FE00; CJK UNIFIED IDEOGRAPH-5292;
52B5 FE00; CJK UNIFIED IDEOGRAPH-5238;
52FB FE00; CJK UNIFIED IDEOGRAPH-5300;
5300 FE00; CJK UNIFIED IDEOGRAPH-52FB;
5355 FE00; CJK UNIFIED IDEOGRAPH-5358;
5358 FE00; CJK UNIFIED IDEOGRAPH-5355;
5373 FE01; CJK UNIFIED IDEOGRAPH-537D;
5377 FE00; CJK UNIFIED IDEOGRAPH-5DFB;
537D FE01; CJK UNIFIED IDEOGRAPH-5373;
53C1 FE00; CJK UNIFIED IDEOGRAPH-53C2;
53C2 FE00; CJK UNIFIED IDEOGRAPH-53C1;
53C3 FE01; CJK UNIFIED IDEOGRAPH-53C4;
53C4 FE00; CJK UNIFIED IDEOGRAPH-53C3;
5415 FE00; CJK UNIFIED IDEOGRAPH-5442;
541E FE00; CJK UNIFIED IDEOGRAPH-5451;
5433 FE00; CJK UNIFIED IDEOGRAPH-5434;
5433 FE01; CJK UNIFIED IDEOGRAPH-5449;
5434 FE00; CJK UNIFIED IDEOGRAPH-5433;
5434 FE01; CJK UNIFIED IDEOGRAPH-5449;
5436 FE00; CJK UNIFIED IDEOGRAPH-5450;
543F FE00; CJK UNIFIED IDEOGRAPH-544A;
5442 FE01; CJK UNIFIED IDEOGRAPH-5415;
5449 FE00; CJK UNIFIED IDEOGRAPH-5433;
5449 FE01; CJK UNIFIED IDEOGRAPH-5434;
544A FE00; CJK UNIFIED IDEOGRAPH-543F;
5450 FE00; CJK UNIFIED IDEOGRAPH-5436;
5451 FE00; CJK UNIFIED IDEOGRAPH-541E;
5527 FE00; CJK UNIFIED IDEOGRAPH-559E;
559E FE00; CJK UNIFIED IDEOGRAPH-5527;
55A9 FE00; CJK UNIFIED IDEOGRAPH-55BB;
55BB FE00; CJK UNIFIED IDEOGRAPH-55A9;
5618 FE00; CJK UNIFIED IDEOGRAPH-5653;
5653 FE00; CJK UNIFIED IDEOGRAPH-5618;
568F FE00; CJK UNIFIED IDEOGRAPH-5694;
5694 FE00; CJK UNIFIED IDEOGRAPH-568F;
56EF FE00; CJK UNIFIED IDEOGRAPH-56FD;

56FD FE00; CJK UNIFIED IDEOGRAPH-56EF;
5708 FE00; CJK UNIFIED IDEOGRAPH-570F;
570E FE00; CJK UNIFIED IDEOGRAPH-5713;
570F FE00; CJK UNIFIED IDEOGRAPH-5708;
5713 FE00; CJK UNIFIED IDEOGRAPH-570E;
5716 FE01; CJK UNIFIED IDEOGRAPH-5717;
5717 FE01; CJK UNIFIED IDEOGRAPH-5716;
5759 FE00; CJK UNIFIED IDEOGRAPH-5DE0;
57D2 FE00; CJK UNIFIED IDEOGRAPH-57D3;
57D3 FE00; CJK UNIFIED IDEOGRAPH-57D2;
5848 FE00; CJK UNIFIED IDEOGRAPH-588D;
5861 FE00; CJK UNIFIED IDEOGRAPH-586B;
586B FE00; CJK UNIFIED IDEOGRAPH-5861;
588D FE00; CJK UNIFIED IDEOGRAPH-5848;
5897 FE00; CJK UNIFIED IDEOGRAPH-589E;
589E FE00; CJK UNIFIED IDEOGRAPH-5897;
58EE FE01; CJK UNIFIED IDEOGRAPH-58EF;
58EF FE00; CJK UNIFIED IDEOGRAPH-58EE;
58FD FE00; CJK UNIFIED IDEOGRAPH-5900;
5900 FE00; CJK UNIFIED IDEOGRAPH-58FD;
5910 FE00; CJK UNIFIED IDEOGRAPH-657B;
5932 FE00; CJK UNIFIED IDEOGRAPH-672C;
5965 FE00; CJK UNIFIED IDEOGRAPH-5967;
5967 FE00; CJK UNIFIED IDEOGRAPH-5965;
5968 FE00; CJK UNIFIED IDEOGRAPH-596C;
5968 FE01; CJK UNIFIED IDEOGRAPH-734E;
596C FE00; CJK UNIFIED IDEOGRAPH-5968;
596C FE01; CJK UNIFIED IDEOGRAPH-734E;
5986 FE00; CJK UNIFIED IDEOGRAPH-599D;
598D FE00; CJK UNIFIED IDEOGRAPH-59F8;
599D FE00; CJK UNIFIED IDEOGRAPH-5986;
59CD FE00; CJK UNIFIED IDEOGRAPH-59D7;
59D7 FE00; CJK UNIFIED IDEOGRAPH-59CD;
59EB FE00; CJK UNIFIED IDEOGRAPH-59EC;
59EC FE01; CJK UNIFIED IDEOGRAPH-59EB;
59F8 FE00; CJK UNIFIED IDEOGRAPH-598D;
5A1B FE01; CJK UNIFIED IDEOGRAPH-5A31;
5A2F FE00; CJK UNIFIED IDEOGRAPH-5A1B;
5A2F FE01; CJK UNIFIED IDEOGRAPH-5A31;
5A31 FE00; CJK UNIFIED IDEOGRAPH-5A1B;
5A31 FE01; CJK UNIFIED IDEOGRAPH-5A2F;
5A55 FE00; CJK UNIFIED IDEOGRAPH-5AAB;
5A7E FE00; CJK UNIFIED IDEOGRAPH-5AAE;
5AAA FE00; CJK UNIFIED IDEOGRAPH-5ABC;

5AAB FE00; CJK UNIFIED IDEOGRAPH-5A55;
5AAE FE00; CJK UNIFIED IDEOGRAPH-5A7E;
5AAF FE00; CJK UNIFIED IDEOGRAPH-5B00;
5ABC FE00; CJK UNIFIED IDEOGRAPH-5AAA;
5B00 FE00; CJK UNIFIED IDEOGRAPH-5AAF;
5B0E FE00; CJK UNIFIED IDEOGRAPH-5B14;
5B14 FE00; CJK UNIFIED IDEOGRAPH-5B0E;
5B24 FE00; CJK UNIFIED IDEOGRAPH-5B37;
5B37 FE00; CJK UNIFIED IDEOGRAPH-5B24;
5B73 FE00; CJK UNIFIED IDEOGRAPH-5B76;
5B76 FE00; CJK UNIFIED IDEOGRAPH-5B73;
5BAB FE00; CJK UNIFIED IDEOGRAPH-5BAE;
5BAE FE00; CJK UNIFIED IDEOGRAPH-5BAB;
5BDB FE00; CJK UNIFIED IDEOGRAPH-5BEC;
5BDC FE00; CJK UNIFIED IDEOGRAPH-5BE7;
5BDD FE00; CJK UNIFIED IDEOGRAPH-5BE2;
5BE2 FE00; CJK UNIFIED IDEOGRAPH-5BDD;
5BE7 FE03; CJK UNIFIED IDEOGRAPH-5BDC;
5BEC FE00; CJK UNIFIED IDEOGRAPH-5BDB;
5C02 FE00; CJK UNIFIED IDEOGRAPH-5C08;
5C06 FE01; CJK UNIFIED IDEOGRAPH-5C07;
5C07 FE00; CJK UNIFIED IDEOGRAPH-5C06;
5C08 FE00; CJK UNIFIED IDEOGRAPH-5C02;
5C13 FE00; CJK UNIFIED IDEOGRAPH-5C14;
5C14 FE00; CJK UNIFIED IDEOGRAPH-5C13;
5C19 FE00; CJK UNIFIED IDEOGRAPH-5C1A;
5C1A FE00; CJK UNIFIED IDEOGRAPH-5C19;
5C2A FE00; CJK UNIFIED IDEOGRAPH-5C2B;
5C2B FE00; CJK UNIFIED IDEOGRAPH-5C2A;
5C36 FE00; CJK UNIFIED IDEOGRAPH-5C37;
5C37 FE00; CJK UNIFIED IDEOGRAPH-5C36;
5C4F FE00; CJK UNIFIED IDEOGRAPH-5C5B;
5C5B FE00; CJK UNIFIED IDEOGRAPH-5C4F;
5CE5 FE00; CJK UNIFIED IDEOGRAPH-5D22;
5D22 FE00; CJK UNIFIED IDEOGRAPH-5CE5;
5DD3 FE00; CJK UNIFIED IDEOGRAPH-5DD4;
5DD4 FE00; CJK UNIFIED IDEOGRAPH-5DD3;
5DE0 FE00; CJK UNIFIED IDEOGRAPH-5759;
5DFB FE00; CJK UNIFIED IDEOGRAPH-5377;
5E21 FE00; CJK UNIFIED IDEOGRAPH-5E32;
5E2F FE00; CJK UNIFIED IDEOGRAPH-5E36;
5E32 FE00; CJK UNIFIED IDEOGRAPH-5E21;
5E36 FE00; CJK UNIFIED IDEOGRAPH-5E2F;
5E76 FE00; CJK UNIFIED IDEOGRAPH-5E77;

5E77 FE00; CJK UNIFIED IDEOGRAPH-5E76;
5E7A FE00; CJK UNIFIED IDEOGRAPH-4E48;
5EC4 FE00; CJK UNIFIED IDEOGRAPH-5ECF;
5ECF FE00; CJK UNIFIED IDEOGRAPH-5EC4;
5F11 FE00; CJK UNIFIED IDEOGRAPH-5F12;
5F12 FE00; CJK UNIFIED IDEOGRAPH-5F11;
5F37 FE00; CJK UNIFIED IDEOGRAPH-5F3A;
5F39 FE00; CJK UNIFIED IDEOGRAPH-5F3E;
5F3A FE00; CJK UNIFIED IDEOGRAPH-5F37;
5F3E FE00; CJK UNIFIED IDEOGRAPH-5F39;
5F50 FE00; CJK UNIFIED IDEOGRAPH-5F51;
5F51 FE00; CJK UNIFIED IDEOGRAPH-5F50;
5F54 FE00; CJK UNIFIED IDEOGRAPH-5F55;
5F55 FE00; CJK UNIFIED IDEOGRAPH-5F54;
5F59 FE00; CJK UNIFIED IDEOGRAPH-5F5A;
5F5A FE00; CJK UNIFIED IDEOGRAPH-5F59;
5F5B FE00; CJK UNIFIED IDEOGRAPH-5F5C;
5F5C FE00; CJK UNIFIED IDEOGRAPH-5F5B;
5F5D FE00; CJK UNIFIED IDEOGRAPH-5F5E;
5F5E FE00; CJK UNIFIED IDEOGRAPH-5F5D;
5F65 FE00; CJK UNIFIED IDEOGRAPH-5F66;
5F66 FE00; CJK UNIFIED IDEOGRAPH-5F65;
5FB3 FE00; CJK UNIFIED IDEOGRAPH-5FB7;
5FB4 FE00; CJK UNIFIED IDEOGRAPH-5FB5;
5FB5 FE00; CJK UNIFIED IDEOGRAPH-5FB4;
5FB7 FE00; CJK UNIFIED IDEOGRAPH-5FB3;
6075 FE01; CJK UNIFIED IDEOGRAPH-60E0;
6085 FE00; CJK UNIFIED IDEOGRAPH-60A6;
609E FE00; CJK UNIFIED IDEOGRAPH-60AE;
60A6 FE00; CJK UNIFIED IDEOGRAPH-6085;
60AE FE00; CJK UNIFIED IDEOGRAPH-609E;
60B3 FE00; CJK UNIFIED IDEOGRAPH-60EA;
60E0 FE00; CJK UNIFIED IDEOGRAPH-6075;
60EA FE00; CJK UNIFIED IDEOGRAPH-60B3;
6120 FE00; CJK UNIFIED IDEOGRAPH-614D;
613C FE00; CJK UNIFIED IDEOGRAPH-614E;
614D FE00; CJK UNIFIED IDEOGRAPH-6120;
614E FE02; CJK UNIFIED IDEOGRAPH-613C;
6229 FE00; CJK UNIFIED IDEOGRAPH-622C;
622C FE00; CJK UNIFIED IDEOGRAPH-6229;
622F FE00; CJK UNIFIED IDEOGRAPH-6231;
6231 FE00; CJK UNIFIED IDEOGRAPH-622F;
6236 FE00; CJK UNIFIED IDEOGRAPH-6237;
6236 FE01; CJK UNIFIED IDEOGRAPH-6238;

6237 FE00; CJK UNIFIED IDEOGRAPH-6236;
6237 FE01; CJK UNIFIED IDEOGRAPH-6238;
6238 FE00; CJK UNIFIED IDEOGRAPH-6236;
6238 FE01; CJK UNIFIED IDEOGRAPH-6237;
623B FE00; CJK UNIFIED IDEOGRAPH-623E;
623E FE00; CJK UNIFIED IDEOGRAPH-623B;
629B FE00; CJK UNIFIED IDEOGRAPH-62CB;
629C FE00; CJK UNIFIED IDEOGRAPH-62D4;
62CB FE00; CJK UNIFIED IDEOGRAPH-629B;
62D4 FE01; CJK UNIFIED IDEOGRAPH-629C;
6329 FE00; CJK UNIFIED IDEOGRAPH-635D;
633F FE00; CJK UNIFIED IDEOGRAPH-63D2;
633F FE01; CJK UNIFIED IDEOGRAPH-63F7;
634F FE00; CJK UNIFIED IDEOGRAPH-63D1;
635C FE00; CJK UNIFIED IDEOGRAPH-641C;
635D FE00; CJK UNIFIED IDEOGRAPH-6329;
63B2 FE00; CJK UNIFIED IDEOGRAPH-63ED;
63D1 FE00; CJK UNIFIED IDEOGRAPH-634F;
63D2 FE00; CJK UNIFIED IDEOGRAPH-633F;
63D2 FE01; CJK UNIFIED IDEOGRAPH-63F7;
63ED FE00; CJK UNIFIED IDEOGRAPH-63B2;
63F7 FE00; CJK UNIFIED IDEOGRAPH-633F;
63F7 FE01; CJK UNIFIED IDEOGRAPH-63D2;
63FA FE00; CJK UNIFIED IDEOGRAPH-6416;
63FA FE01; CJK UNIFIED IDEOGRAPH-6447;
63FE FE00; CJK UNIFIED IDEOGRAPH-6435;
6416 FE00; CJK UNIFIED IDEOGRAPH-63FA;
6416 FE01; CJK UNIFIED IDEOGRAPH-6447;
641C FE01; CJK UNIFIED IDEOGRAPH-635C;
6435 FE00; CJK UNIFIED IDEOGRAPH-63FE;
6447 FE00; CJK UNIFIED IDEOGRAPH-63FA;
6447 FE01; CJK UNIFIED IDEOGRAPH-6416;
6483 FE00; CJK UNIFIED IDEOGRAPH-64CA;
64CA FE00; CJK UNIFIED IDEOGRAPH-6483;
654E FE00; CJK UNIFIED IDEOGRAPH-6559;
6553 FE00; CJK UNIFIED IDEOGRAPH-655A;
6559 FE00; CJK UNIFIED IDEOGRAPH-654E;
655A FE00; CJK UNIFIED IDEOGRAPH-6553;
657B FE00; CJK UNIFIED IDEOGRAPH-5910;
65E2 FE01; CJK UNIFIED IDEOGRAPH-65E3;
65E3 FE01; CJK UNIFIED IDEOGRAPH-65E2;
6602 FE00; CJK UNIFIED IDEOGRAPH-663B;
663B FE00; CJK UNIFIED IDEOGRAPH-6602;
665A FE00; CJK UNIFIED IDEOGRAPH-6669;

6669 FE00; CJK UNIFIED IDEOGRAPH-665A;
66A8 FE00; CJK UNIFIED IDEOGRAPH-66C1;
66C1 FE00; CJK UNIFIED IDEOGRAPH-66A8;
66FD FE00; CJK UNIFIED IDEOGRAPH-66FE;
66FE FE00; CJK UNIFIED IDEOGRAPH-66FD;
672C FE00; CJK UNIFIED IDEOGRAPH-5932;
67B4 FE00; CJK UNIFIED IDEOGRAPH-67FA;
67E5 FE00; CJK UNIFIED IDEOGRAPH-67FB;
67F5 FE00; CJK UNIFIED IDEOGRAPH-6805;
67FA FE01; CJK UNIFIED IDEOGRAPH-67B4;
67FB FE00; CJK UNIFIED IDEOGRAPH-67E5;
6805 FE00; CJK UNIFIED IDEOGRAPH-67F5;
68B2 FE00; CJK UNIFIED IDEOGRAPH-68C1;
68C1 FE00; CJK UNIFIED IDEOGRAPH-68B2;
6961 FE00; CJK UNIFIED IDEOGRAPH-6986;
6982 FE00; CJK UNIFIED IDEOGRAPH-69EA;
6985 FE00; CJK UNIFIED IDEOGRAPH-69B2;
6986 FE00; CJK UNIFIED IDEOGRAPH-6961;
699D FE00; CJK UNIFIED IDEOGRAPH-6A27;
69B2 FE00; CJK UNIFIED IDEOGRAPH-6985;
69C7 FE00; CJK UNIFIED IDEOGRAPH-69D9;
69D8 FE00; CJK UNIFIED IDEOGRAPH-6A23;
69D9 FE00; CJK UNIFIED IDEOGRAPH-69C7;
69EA FE01; CJK UNIFIED IDEOGRAPH-6982;
6A23 FE00; CJK UNIFIED IDEOGRAPH-69D8;
6A27 FE00; CJK UNIFIED IDEOGRAPH-699D;
6A2A FE00; CJK UNIFIED IDEOGRAPH-6A6B;
6A6B FE00; CJK UNIFIED IDEOGRAPH-6A2A;
6B65 FE00; CJK UNIFIED IDEOGRAPH-6B69;
6B69 FE00; CJK UNIFIED IDEOGRAPH-6B65;
6B72 FE01; CJK UNIFIED IDEOGRAPH-6B73;
6B73 FE00; CJK UNIFIED IDEOGRAPH-6B72;
6B7F FE00; CJK UNIFIED IDEOGRAPH-6B81;
6B81 FE00; CJK UNIFIED IDEOGRAPH-6B7F;
6BBB FE01; CJK UNIFIED IDEOGRAPH-6BBC;
6BBC FE00; CJK UNIFIED IDEOGRAPH-6BBB;
6BC0 FE00; CJK UNIFIED IDEOGRAPH-6BC1;
6BC1 FE00; CJK UNIFIED IDEOGRAPH-6BC0;
6BCE FE00; CJK UNIFIED IDEOGRAPH-6BCF;
6BCF FE00; CJK UNIFIED IDEOGRAPH-6BCE;
6C32 FE00; CJK UNIFIED IDEOGRAPH-6C33;
6C33 FE00; CJK UNIFIED IDEOGRAPH-6C32;
6C5A FE00; CJK UNIFIED IDEOGRAPH-6C61;
6C61 FE00; CJK UNIFIED IDEOGRAPH-6C5A;

6C92 FE00; CJK UNIFIED IDEOGRAPH-6CA1;
6CA1 FE00; CJK UNIFIED IDEOGRAPH-6C92;
6D44 FE00; CJK UNIFIED IDEOGRAPH-6DE8;
6D89 FE00; CJK UNIFIED IDEOGRAPH-6E09;
6D97 FE00; CJK UNIFIED IDEOGRAPH-6D9A;
6D99 FE00; CJK UNIFIED IDEOGRAPH-6DDA;
6D9A FE00; CJK UNIFIED IDEOGRAPH-6D97;
6DDA FE01; CJK UNIFIED IDEOGRAPH-6D99;
6DE5 FE00; CJK UNIFIED IDEOGRAPH-6E0C;
6DE8 FE00; CJK UNIFIED IDEOGRAPH-6D44;
6DF8 FE00; CJK UNIFIED IDEOGRAPH-6E05;
6E05 FE00; CJK UNIFIED IDEOGRAPH-6DF8;
6E07 FE00; CJK UNIFIED IDEOGRAPH-6E34;
6E09 FE00; CJK UNIFIED IDEOGRAPH-6D89;
6E0C FE00; CJK UNIFIED IDEOGRAPH-6DE5;
6E29 FE00; CJK UNIFIED IDEOGRAPH-6EAB;
6E34 FE00; CJK UNIFIED IDEOGRAPH-6E07;
6E88 FE00; CJK UNIFIED IDEOGRAPH-6F59;
6E89 FE00; CJK UNIFIED IDEOGRAPH-6F11;
6EAB FE00; CJK UNIFIED IDEOGRAPH-6E29;
6EDA FE00; CJK UNIFIED IDEOGRAPH-6EFE;
6EFE FE00; CJK UNIFIED IDEOGRAPH-6EDA;
6F11 FE00; CJK UNIFIED IDEOGRAPH-6E89;
6F59 FE00; CJK UNIFIED IDEOGRAPH-6E88;
6F5B FE00; CJK UNIFIED IDEOGRAPH-6FF3;
6FF3 FE00; CJK UNIFIED IDEOGRAPH-6F5B;
7028 FE00; CJK UNIFIED IDEOGRAPH-702C;
702C FE00; CJK UNIFIED IDEOGRAPH-7028;
70BA FE00; CJK UNIFIED IDEOGRAPH-7232;
712D FE00; CJK UNIFIED IDEOGRAPH-7162;
7155 FE00; CJK UNIFIED IDEOGRAPH-7199;
7162 FE00; CJK UNIFIED IDEOGRAPH-712D;
7174 FE00; CJK UNIFIED IDEOGRAPH-7185;
7185 FE00; CJK UNIFIED IDEOGRAPH-7174;
7199 FE00; CJK UNIFIED IDEOGRAPH-7155;
722D FE00; CJK UNIFIED IDEOGRAPH-4E89;
7232 FE00; CJK UNIFIED IDEOGRAPH-70BA;
72B6 FE00; CJK UNIFIED IDEOGRAPH-72C0;
72C0 FE01; CJK UNIFIED IDEOGRAPH-72B6;
734E FE00; CJK UNIFIED IDEOGRAPH-5968;
734E FE01; CJK UNIFIED IDEOGRAPH-596C;
7464 FE00; CJK UNIFIED IDEOGRAPH-7476;
7476 FE00; CJK UNIFIED IDEOGRAPH-7464;
74F6 FE00; CJK UNIFIED IDEOGRAPH-7501;

7501 FE00; CJK UNIFIED IDEOGRAPH-74F6;
7522 FE00; CJK UNIFIED IDEOGRAPH-7523;
7523 FE00; CJK UNIFIED IDEOGRAPH-7522;
75E9 FE00; CJK UNIFIED IDEOGRAPH-7626;
7626 FE00; CJK UNIFIED IDEOGRAPH-75E9;
76A1 FE00; CJK UNIFIED IDEOGRAPH-76A5;
76A5 FE00; CJK UNIFIED IDEOGRAPH-76A1;
771E FE01; CJK UNIFIED IDEOGRAPH-771F;
771F FE02; CJK UNIFIED IDEOGRAPH-771E;
773E FE00; CJK UNIFIED IDEOGRAPH-8846;
7814 FE00; CJK UNIFIED IDEOGRAPH-784F;
784F FE00; CJK UNIFIED IDEOGRAPH-7814;
797F FE01; CJK UNIFIED IDEOGRAPH-7984;
7984 FE00; CJK UNIFIED IDEOGRAPH-797F;
79BF FE00; CJK UNIFIED IDEOGRAPH-79C3;
79C3 FE00; CJK UNIFIED IDEOGRAPH-79BF;
7A05 FE00; CJK UNIFIED IDEOGRAPH-7A0E;
7A0E FE00; CJK UNIFIED IDEOGRAPH-7A05;
7A42 FE00; CJK UNIFIED IDEOGRAPH-7A57;
7A57 FE00; CJK UNIFIED IDEOGRAPH-7A42;
7B5D FE00; CJK UNIFIED IDEOGRAPH-7B8F;
7B8F FE00; CJK UNIFIED IDEOGRAPH-7B5D;
7BB3 FE00; CJK UNIFIED IDEOGRAPH-7C08;
7BE1 FE00; CJK UNIFIED IDEOGRAPH-7C12;
7C08 FE00; CJK UNIFIED IDEOGRAPH-7BB3;
7C12 FE00; CJK UNIFIED IDEOGRAPH-7BE1;
7CA4 FE00; CJK UNIFIED IDEOGRAPH-7CB5;
7CB5 FE00; CJK UNIFIED IDEOGRAPH-7CA4;
7D55 FE00; CJK UNIFIED IDEOGRAPH-7D76;
7D76 FE00; CJK UNIFIED IDEOGRAPH-7D55;
7DA0 FE01; CJK UNIFIED IDEOGRAPH-7DD1;
7DD1 FE00; CJK UNIFIED IDEOGRAPH-7DA0;
7DD2 FE00; CJK UNIFIED IDEOGRAPH-7DD6;
7DD6 FE00; CJK UNIFIED IDEOGRAPH-7DD2;
7DE3 FE00; CJK UNIFIED IDEOGRAPH-7E01;
7DFC FE00; CJK UNIFIED IDEOGRAPH-7E15;
7E01 FE00; CJK UNIFIED IDEOGRAPH-7DE3;
7E15 FE00; CJK UNIFIED IDEOGRAPH-7DFC;
7E48 FE00; CJK UNIFIED IDEOGRAPH-7E66;
7E66 FE00; CJK UNIFIED IDEOGRAPH-7E48;
7FAE FE00; CJK UNIFIED IDEOGRAPH-7FB9;
7FB9 FE00; CJK UNIFIED IDEOGRAPH-7FAE;
7FF6 FE00; CJK UNIFIED IDEOGRAPH-7FFA;
7FFA FE01; CJK UNIFIED IDEOGRAPH-7FF6;

80FC FE00; CJK UNIFIED IDEOGRAPH-8141;
812B FE00; CJK UNIFIED IDEOGRAPH-8131;
8131 FE00; CJK UNIFIED IDEOGRAPH-812B;
8141 FE00; CJK UNIFIED IDEOGRAPH-80FC;
817D FE00; CJK UNIFIED IDEOGRAPH-8183;
8183 FE00; CJK UNIFIED IDEOGRAPH-817D;
8203 FE00; CJK UNIFIED IDEOGRAPH-8204;
8204 FE01; CJK UNIFIED IDEOGRAPH-8203;
820D FE00; CJK UNIFIED IDEOGRAPH-820E;
820E FE00; CJK UNIFIED IDEOGRAPH-820D;
8216 FE00; CJK UNIFIED IDEOGRAPH-8217;
8217 FE00; CJK UNIFIED IDEOGRAPH-8216;
8358 FE00; CJK UNIFIED IDEOGRAPH-838A;
838A FE00; CJK UNIFIED IDEOGRAPH-8358;
83D1 FE00; CJK UNIFIED IDEOGRAPH-8458;
8458 FE00; CJK UNIFIED IDEOGRAPH-83D1;
8480 FE00; CJK UNIFIED IDEOGRAPH-8495;
848B FE00; CJK UNIFIED IDEOGRAPH-8523;
848D FE00; CJK UNIFIED IDEOGRAPH-853F;
8495 FE00; CJK UNIFIED IDEOGRAPH-8480;
8523 FE00; CJK UNIFIED IDEOGRAPH-848B;
853F FE00; CJK UNIFIED IDEOGRAPH-848D;
8570 FE00; CJK UNIFIED IDEOGRAPH-8580;
8580 FE00; CJK UNIFIED IDEOGRAPH-8570;
85AB FE00; CJK UNIFIED IDEOGRAPH-85B0;
85B0 FE00; CJK UNIFIED IDEOGRAPH-85AB;
85F4 FE00; CJK UNIFIED IDEOGRAPH-860A;
860A FE00; CJK UNIFIED IDEOGRAPH-85F4;
865A FE00; CJK UNIFIED IDEOGRAPH-865B;
865B FE00; CJK UNIFIED IDEOGRAPH-865A;
86FB FE00; CJK UNIFIED IDEOGRAPH-8715;
8715 FE00; CJK UNIFIED IDEOGRAPH-86FB;
8846 FE00; CJK UNIFIED IDEOGRAPH-773E;
885B FE00; CJK UNIFIED IDEOGRAPH-885E;
885E FE00; CJK UNIFIED IDEOGRAPH-885B;
886E FE00; CJK UNIFIED IDEOGRAPH-889E;
889E FE00; CJK UNIFIED IDEOGRAPH-886E;
88C5 FE00; CJK UNIFIED IDEOGRAPH-88DD;
88DD FE00; CJK UNIFIED IDEOGRAPH-88C5;
8A2E FE00; CJK UNIFIED IDEOGRAPH-8A7D;
8A7D FE00; CJK UNIFIED IDEOGRAPH-8A2E;
8AAA FE02; CJK UNIFIED IDEOGRAPH-8AAC;
8AAC FE00; CJK UNIFIED IDEOGRAPH-8AAA;
8ACC FE00; CJK UNIFIED IDEOGRAPH-8AEB;

8AEB FE00; CJK UNIFIED IDEOGRAPH-8ACC;
8B20 FE00; CJK UNIFIED IDEOGRAPH-8B21;
8B21 FE00; CJK UNIFIED IDEOGRAPH-8B20;
8C5C FE00; CJK UNIFIED IDEOGRAPH-8C63;
8C63 FE00; CJK UNIFIED IDEOGRAPH-8C5C;
8D70 FE00; CJK UNIFIED IDEOGRAPH-8D71;
8D71 FE00; CJK UNIFIED IDEOGRAPH-8D70;
8EFF FE00; CJK UNIFIED IDEOGRAPH-8F27;
8F1C FE00; CJK UNIFIED IDEOGRAPH-8F3A;
8F27 FE00; CJK UNIFIED IDEOGRAPH-8EFF;
8F3A FE00; CJK UNIFIED IDEOGRAPH-8F1C;
8F3C FE00; CJK UNIFIED IDEOGRAPH-8F40;
8F40 FE00; CJK UNIFIED IDEOGRAPH-8F3C;
8FBE FE00; CJK UNIFIED IDEOGRAPH-8FD6;
8FD6 FE00; CJK UNIFIED IDEOGRAPH-8FBE;
8FF8 FE00; CJK UNIFIED IDEOGRAPH-902C;
902C FE00; CJK UNIFIED IDEOGRAPH-8FF8;
9059 FE00; CJK UNIFIED IDEOGRAPH-9065;
9065 FE00; CJK UNIFIED IDEOGRAPH-9059;
90A2 FE00; CJK UNIFIED IDEOGRAPH-90C9;
90C9 FE00; CJK UNIFIED IDEOGRAPH-90A2;
90CE FE01; CJK UNIFIED IDEOGRAPH-90DE;
90DE FE01; CJK UNIFIED IDEOGRAPH-90CE;
90F7 FE00; CJK UNIFIED IDEOGRAPH-9109;
90F7 FE01; CJK UNIFIED IDEOGRAPH-9115;
9109 FE00; CJK UNIFIED IDEOGRAPH-90F7;
9109 FE01; CJK UNIFIED IDEOGRAPH-9115;
9115 FE00; CJK UNIFIED IDEOGRAPH-90F7;
9115 FE01; CJK UNIFIED IDEOGRAPH-9109;
9196 FE00; CJK UNIFIED IDEOGRAPH-919E;
919E FE00; CJK UNIFIED IDEOGRAPH-9196;
91A4 FE00; CJK UNIFIED IDEOGRAPH-91AC;
91AC FE00; CJK UNIFIED IDEOGRAPH-91A4;
9203 FE00; CJK UNIFIED IDEOGRAPH-9292;
9292 FE00; CJK UNIFIED IDEOGRAPH-9203;
92B3 FE00; CJK UNIFIED IDEOGRAPH-92ED;
92ED FE00; CJK UNIFIED IDEOGRAPH-92B3;
9304 FE01; CJK UNIFIED IDEOGRAPH-9332;
932C FE00; CJK UNIFIED IDEOGRAPH-934A;
9332 FE00; CJK UNIFIED IDEOGRAPH-9304;
934A FE01; CJK UNIFIED IDEOGRAPH-932C;
93AD FE00; CJK UNIFIED IDEOGRAPH-93AE;
93AE FE00; CJK UNIFIED IDEOGRAPH-93AD;
95B1 FE00; CJK UNIFIED IDEOGRAPH-95B2;

95B2 FE00; CJK UNIFIED IDEOGRAPH-95B1;
9667 FE00; CJK UNIFIED IDEOGRAPH-9689;
9689 FE00; CJK UNIFIED IDEOGRAPH-9667;
9751 FE00; CJK UNIFIED IDEOGRAPH-9752;
9752 FE00; CJK UNIFIED IDEOGRAPH-9751;
9759 FE00; CJK UNIFIED IDEOGRAPH-975C;
975C FE00; CJK UNIFIED IDEOGRAPH-9759;
976D FE00; CJK UNIFIED IDEOGRAPH-9771;
9771 FE00; CJK UNIFIED IDEOGRAPH-976D;
9839 FE00; CJK UNIFIED IDEOGRAPH-983D;
983D FE00; CJK UNIFIED IDEOGRAPH-9839;
984F FE00; CJK UNIFIED IDEOGRAPH-9854;
9854 FE00; CJK UNIFIED IDEOGRAPH-984F;
985A FE00; CJK UNIFIED IDEOGRAPH-985B;
985B FE00; CJK UNIFIED IDEOGRAPH-985A;
98EE FE00; CJK UNIFIED IDEOGRAPH-98F2;
98F2 FE00; CJK UNIFIED IDEOGRAPH-98EE;
9905 FE00; CJK UNIFIED IDEOGRAPH-9920;
9920 FE00; CJK UNIFIED IDEOGRAPH-9905;
99B1 FE00; CJK UNIFIED IDEOGRAPH-99C4;
99C4 FE00; CJK UNIFIED IDEOGRAPH-99B1;
99E2 FE00; CJK UNIFIED IDEOGRAPH-9A08;
9A08 FE00; CJK UNIFIED IDEOGRAPH-99E2;
9AA9 FE00; CJK UNIFIED IDEOGRAPH-9AAB;
9AAB FE00; CJK UNIFIED IDEOGRAPH-9AA9;
9AD8 FE00; CJK UNIFIED IDEOGRAPH-9AD9;
9AD9 FE00; CJK UNIFIED IDEOGRAPH-9AD8;
9AEA FE00; CJK UNIFIED IDEOGRAPH-9AEE;
9AEE FE00; CJK UNIFIED IDEOGRAPH-9AEA;
9B2C FE00; CJK UNIFIED IDEOGRAPH-9B2D;
9B2D FE00; CJK UNIFIED IDEOGRAPH-9B2C;
9C1B FE00; CJK UNIFIED IDEOGRAPH-9C2E;
9C2E FE00; CJK UNIFIED IDEOGRAPH-9C1B;
9CEF FE00; CJK UNIFIED IDEOGRAPH-9CF3;
9CF3 FE00; CJK UNIFIED IDEOGRAPH-9CEF;
9D87 FE00; CJK UNIFIED IDEOGRAPH-9DAB;
9DAB FE00; CJK UNIFIED IDEOGRAPH-9D87;
9DC6 FE00; CJK UNIFIED IDEOGRAPH-9DCF;
9DCF FE00; CJK UNIFIED IDEOGRAPH-9DC6;
9EAA FE00; CJK UNIFIED IDEOGRAPH-9EAB;
9EAB FE00; CJK UNIFIED IDEOGRAPH-9EAA;
9EBC FE00; CJK UNIFIED IDEOGRAPH-9EBD;
9EBD FE00; CJK UNIFIED IDEOGRAPH-9EBC;
9EC3 FE00; CJK UNIFIED IDEOGRAPH-9EC4;

9EC4 FE00; CJK UNIFIED IDEOGRAPH-9EC3;
9ED1 FE00; CJK UNIFIED IDEOGRAPH-9ED2;
9ED2 FE00; CJK UNIFIED IDEOGRAPH-9ED1;
EOF