

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N2538

Date: 2022-03-13

Source:	China
Meeting:	IRG#58
Title:	Feedback on IRGN2538 by China
Status:	Member's submission
Actions required:	To be considered by IRG
Distribution:	IRG
Author:	TAO Yang
Medium:	Electronic
Pages:	3
Appendixes:	0

Part 1 Traceability of difference causes

It is a good way to explore this problem from the original source.

As it shows in GB 18030-2005 信息技术 中文编码字符集, U+5F50 is the radical character of 彗/彗/彗. This relationship can be seen more clearly in the code chart.

E5	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
A		灑 6FC9	灑 6FA7	灑 6FB9	灑 6FB6	灑 6FC2	灑 6FE1	灑 6FEE	灑 6FDE	灑 6FE0	灑 6FEP	灑 701A	灑 7023	灑 701B	灑 7039	灑 7035
B	灑 704F	灑 705E	灑 5B80	宀 5B84	宀 5B95	宀 5B93	宀 5BA5	宀 5BB8	宀 752F	宀 9A9E	宀 6434	宀 5BE4	宀 5BEE	宀 8930	宀 5BF0	宀 8E47
C	灑 8B07	灑 8FB6	灑 8FD3	灑 8FD5	灑 8FE5	灑 8FEE	灑 8FE4	灑 8FE9	灑 8FE6	灑 8FF3	灑 8FE8	灑 9005	灑 9004	灑 900B	灑 9026	灑 9011
D	灑 900D	灑 9016	灑 9021	灑 9035	灑 9036	灑 902D	灑 902F	灑 9044	灑 9051	灑 9052	灑 9050	灑 9068	灑 9058	灑 9062	灑 905B	灑 66B9
E	灑 9074	灑 907D	灑 9082	灑 9088	灑 9083	灑 908B	彗 5F50	彗 5F57	彗 5F56	彗 5F58	彗 5C3B	彗 54AB	彗 5C50	彗 5C59	彗 5B71	彗 5C63
F	彗 5C66	彗 7FBC	彗 5P2A	彗 5F29	彗 5F2D	彗 8274	彗 5F3C	彗 9B3B	彗 5C6E	彗 5981	彗 5983	彗 598D	彗 59A9	彗 59AA	彗 59A3	

GB 18030-2005 信息技术 中文编码字符集

5F50 彗 G0-6986 H-0800 T3-2140 J1-3C77 K2-313A V1-9454	彗 彗 彗 彗 彗	5F58 彗 G0-6988 H2-08F4 T3-3888 J3-743C K1-8F9E	彗 彗 彗 彗 彗
5F51 彗 G0-2950 H-88CA T3-2141 G0-6780 K2-313B	彗 彗 彗 彗 彗	5F59 彗 G1-7820 H2-0784 T1-4310 J0-5741 K2-1044 G2-8873	彗 彗 彗 彗 彗
5F52 彗 G0-3969 T3-224A	彗 彗	5F5A 彗 G0-2952 T3-4102	彗 彗
5F53 彗 G0-3031 T3-2281 J0-4576 K2-313C	彗 彗 彗 彗 彗	5F5B 彗 G0-2947 H2-0784 J0-5741 K2-1044 G2-8873	彗 彗 彗 彗 彗
5F54 彗 G0-2951 H2-CC74 T2-2655 J1-3-743B K1-8E79	彗 彗 彗 彗 彗	5F5C 彗 G0-2947 H2-0784 J0-5741 K2-1044 G2-8873	彗 彗 彗 彗 彗
5F55 彗 G0-423C T3-2071	彗 彗	5F5D 彗 G0-2948 H2-CC55 T3-0784 J0-5741 K1-8A41	彗 彗 彗 彗 彗
5F56 彗 G0-6988 H2-CE80 T2-2A3E J0-5741 K0-0327	彗 彗 彗 彗 彗	5F5E 彗 G0-2944 T1-7641	彗 彗
5F57 彗 G0-6287 H2-818B T1-5A31 J0-5742 K0-7832 V0-8872	彗 彗 彗 彗 彗	5F5F 彗 G0-2944 G0-2945 G0-2946	彗 彗 彗 彗 彗
		5F60 彗 G0-2944 T2-0784 J1-3C77 K2-313E	彗 彗 彗 彗

Code chart of URO

Obviously, only the glyph from G-source is different from other sources. This difference needs to

find an explanation path from the development history of dictionaries.

According to the timeline, the evolution of this radical has also become clear.

Dictionary Name	Radical in Checklists	Head characters																																																						
正字通 ZhengZiTong																																																								
康熙字典 KangXi Zidian																																																								
中华大字典 Zhonghua Dazidian																																																								
新华字典 Xinhua Zidian		<table border="1"> <tr> <td>(52)</td> <td>当郎 1128</td> </tr> <tr> <td>彣部</td> <td>录 2175</td> </tr> <tr> <td>(彣)</td> <td>彣 3025</td> </tr> <tr> <td></td> <td>彣 4469</td> </tr> <tr> <td></td> <td>(彣)</td> </tr> <tr> <td>归 2503</td> <td>3024</td> </tr> <tr> <td>彣 5024</td> <td>(彣)</td> </tr> <tr> <td>当 1127</td> <td>6164</td> </tr> <tr> <td>1133</td> <td>彣 6164</td> </tr> </table>	(52)	当郎 1128	彣部	录 2175	(彣)	彣 3025		彣 4469		(彣)	归 2503	3024	彣 5024	(彣)	当 1127	6164	1133	彣 6164																																				
(52)	当郎 1128																																																							
彣部	录 2175																																																							
(彣)	彣 3025																																																							
	彣 4469																																																							
	(彣)																																																							
归 2503	3024																																																							
彣 5024	(彣)																																																							
当 1127	6164																																																							
1133	彣 6164																																																							
现代汉语词典 Xiandai Hanyu Cidian		<table border="1"> <tr> <td>50</td> <td></td> </tr> <tr> <td>彣部</td> <td></td> </tr> <tr> <td>归 487</td> <td></td> </tr> <tr> <td>彣° 193</td> <td></td> </tr> <tr> <td>当° 258</td> <td></td> </tr> <tr> <td>261</td> <td></td> </tr> <tr> <td>寻 1483</td> <td></td> </tr> <tr> <td>灵 823</td> <td></td> </tr> <tr> <td>彣 1695</td> <td></td> </tr> <tr> <td>彣 583</td> <td></td> </tr> <tr> <td>(寻) 1483</td> <td></td> </tr> <tr> <td>(寻) 1483</td> <td></td> </tr> <tr> <td>(归)° 487</td> <td></td> </tr> <tr> <td>彣 1606</td> <td></td> </tr> <tr> <td>(彣) 1606</td> <td></td> </tr> </table> <table border="1"> <tr> <td>[50]</td> <td></td> </tr> <tr> <td>彣部</td> <td></td> </tr> <tr> <td>彣 1322</td> <td></td> </tr> <tr> <td>彣 1682</td> <td></td> </tr> <tr> <td>(彣) 580</td> <td></td> </tr> <tr> <td>(彣) 1537</td> <td></td> </tr> <tr> <td>彣 1537</td> <td></td> </tr> <tr> <td>彣° 793</td> <td></td> </tr> <tr> <td>彣 796</td> <td></td> </tr> </table> <table border="1"> <tr> <td>[50]</td> <td></td> </tr> <tr> <td>彣部</td> <td></td> </tr> <tr> <td>录 843</td> <td></td> </tr> </table>	50		彣部		归 487		彣° 193		当° 258		261		寻 1483		灵 823		彣 1695		彣 583		(寻) 1483		(寻) 1483		(归)° 487		彣 1606		(彣) 1606		[50]		彣部		彣 1322		彣 1682		(彣) 580		(彣) 1537		彣 1537		彣° 793		彣 796		[50]		彣部		录 843	
50																																																								
彣部																																																								
归 487																																																								
彣° 193																																																								
当° 258																																																								
261																																																								
寻 1483																																																								
灵 823																																																								
彣 1695																																																								
彣 583																																																								
(寻) 1483																																																								
(寻) 1483																																																								
(归)° 487																																																								
彣 1606																																																								
(彣) 1606																																																								
[50]																																																								
彣部																																																								
彣 1322																																																								
彣 1682																																																								
(彣) 580																																																								
(彣) 1537																																																								
彣 1537																																																								
彣° 793																																																								
彣 796																																																								
[50]																																																								
彣部																																																								
录 843																																																								

According to the principle of merging approximate radicals to facilitate shape near glyphs retrieval,

the characters contain components 𠃉/𠃊/ 𠃋/ 𠃌 are integrated into one dictionary catalogue. The traditional dictionaries determined the main radical as 𠃋, then the glyphs of sources except G are in accordance with traditional regulations. But characters have undergone a long period of reform in mainland China, some of the characters contain 𠃋 are integrated to 𠃉, then the quantity of characters with component 𠃉 become a majority, the modern dictionaries of mainland China accept this adjustment and use 𠃉 as the radical glyph. That's why only the G glyph is different to the others.

Part 2 Suggestions on the use of ID in Macao

Returning to the question raised by the representative of Macao.

The glyph 𠃉 has been used as a person name on ID card. Unify the glyph to U+5F50 or add an UNC, bases on whether the civil affairs system of Macao compiles this glyph on code position 5F50. If yes, it would be better to unify to 5F50 and add a M-source. If no, adding to a new code point is feasible.

Part 3 China's opinion on G-source glyph

Previously, China believed that according to the radical rules of traditional dictionaries, the glyph should be modified to 𠃋 to keep the code chart consistent, and 𠃉 should added as a new character on another code point.

However, this idea does not take into account the risks in practical applications.

The technicians of the Ministry noticed also noticed the proposals on dis-unifying U+5F50 𠃉 (IRGN2414, IRGN2509, IRGN2538, IRGN2543) and thought it would be a great impact for affecting 56,000 person names and 400 place names in the ID system. Thus, keeping it unchanged or a compromise is hoped.

(End of document)