

Subject: Issues in "IICore Radical-Stroke Chart" (IICoreRSIndex.pdf)

Date: 2019-12-02

Author: Michel MARIANI

Description

The document "IICore Radical-Stroke Chart" (https://www.unicode.org/versions/IICoreRSIndex.pdf) traditionally lists the Unihan characters belonging to the IICore set of 9,810 CJK unified ideographs in common usage, sorted according to their KangXi radical number and residual stroke count.

Issues

- 1. The IICore character 黄 U+9EC4 is listed twice in a row under Radical 201 黄 (Yellow) + 0 strokes; although this is technically understandable, it is definitely confusing for the reader: for the same combination of radical number and residual stroke count, characters are usually listed only once, without taking into account any radical variants.
2. The 42 IICore characters which happen to belong to the block CJK Unified Ideographs Extension A (U+3400..U+4DBF) are not correctly ordered by ascending code point value.

Example of issue 1

Under Radical 201 黄 (Yellow) + 0 strokes, the same character 黄 U+9EC4 is listed twice in a row:

Table showing character listings for Radical 201 黄 (Yellow) + 0 strokes. It lists characters like 麻 (200), 黄 (201), and 覺 (9ECC) with their respective code points and stroke counts. The character 黄 U+9EC4 is listed twice in a row.

Examples of issue 2

Under Radical 18 刀 (Knife) + 6 strokes, the character 割 U+34E4 is incorrectly listed between 剂 U+5242 and 剂 U+2070E:

Table showing character listings for Radical 18 刀 (Knife) + 6 strokes. It lists characters like 割 (U+34E4), 剗 (U+2070E), and 剗 (U+34E4) with their respective code points and stroke counts. The character 割 U+34E4 is incorrectly listed between 剂 U+5242 and 剂 U+2070E.

Under Radical 196 鳥 (Bird) + 13 strokes, the character 鸚 U+4D08 is incorrectly listed between 鸚 U+9DFF and 鸚 U+9E6D:

Table showing character listings for Radical 196 鳥 (Bird) + 13 strokes. It lists characters like 鸚 (U+4D08), 鸚 (U+9DFF), and 鸚 (U+9E6D) with their respective code points and stroke counts. The character 鸚 U+4D08 is incorrectly listed between 鸚 U+9DFF and 鸚 U+9E6D.

References

List of the 3 Unihan characters with multiple `kRSUnicode` tag values, but with the same combination of radical number and residual stroke count (only the third one belongs to the IICore set):

𠂇	U+4E2C	kRSUnicode	90.0 90'.0
𠂈	U+8FB6	kRSUnicode	162.0 162'.0
黄	U+9EC4	kRSUnicode	201.0 201'.0

Unihan characters belonging to the IICore set of CJK unified ideographs in common usage can only be found in three Unicode blocks:

Block	Range	IICore Set
CJK Unified Ideographs	U+4E00..U+9FFF	9,706
CJK Unified Ideographs Extension A	U+3400..U+4DBF	42
CJK Unified Ideographs Extension B	U+20000..U+2A6DF	62
		9,810

List of the 42 incorrectly ordered Unihan characters belonging to the IICore set, part of the block **CJK Unified Ideographs Extension A**:

刮	U+34E4	kRSUnicode	18.6
𠂇	U+3577	kRSUnicode	30.5
𠂈	U+35A1	kRSUnicode	30.8
𠂉	U+35AD	kRSUnicode	30.8
𠂊	U+35BF	kRSUnicode	30.9
𠂋	U+35CE	kRSUnicode	30.9
𠂌	U+35F3	kRSUnicode	30.12
𠂍	U+35FE	kRSUnicode	30.13
𠂎	U+3960	kRSUnicode	61.9
𠂏	U+39DF	kRSUnicode	64.5
𠂐	U+39F8	kRSUnicode	64.7
𠂑	U+39FE	kRSUnicode	64.8
𠂒	U+3A18	kRSUnicode	64.9
𠂓	U+3A52	kRSUnicode	64.13
𠂔	U+3A67	kRSUnicode	64.15
𠂕	U+3B39	kRSUnicode	130.8
𠂖	U+3B4E	kRSUnicode	75.4
𠂗	U+3DE7	kRSUnicode	86.10
𠂘	U+3DEB	kRSUnicode	86.11
𠂙	U+3E74	kRSUnicode	94.7
𠂚	U+3ED0	kRSUnicode	96.7
𠂛	U+4065	kRSUnicode	109.11
𠂜	U+406A	kRSUnicode	109.11
𠂝	U+40BB	kRSUnicode	112.8
𠂞	U+40DF	kRSUnicode	112.12
𠂟	U+4137	kRSUnicode	115.5
𠂠	U+44EA	kRSUnicode	140.8
𠂡	U+4606	kRSUnicode	142.16
𠂢	U+4723	kRSUnicode	149'.4
𠂣	U+47F4	kRSUnicode	157.7
𠂤	U+48B5	kRSUnicode	163.4
𠂥	U+48BC	kRSUnicode	163.6
𠂦	U+48C5	kRSUnicode	163.7
𠂧	U+48D3	kRSUnicode	163.10
𠂨	U+49D1	kRSUnicode	170.8
𠂩	U+4A12	kRSUnicode	173.6
𠂪	U+4AB8	kRSUnicode	181.5
𠂫	U+4C7D	kRSUnicode	195.10
𠂬	U+4C81	kRSUnicode	195.11
𠂭	U+4C85	kRSUnicode	195.11
𠂮	U+4CB3	kRSUnicode	196.4
𠂯	U+4D08	kRSUnicode	196.13

Notes

- For what it's worth, and in case it may be useful, the attached JSON file **IICoreRS.json** provides an issue-free basic list of all Unihan characters belonging to the IICore set, sorted by KangXi radical number and residual stroke count.
- The recently released document "Unihan Core 2020 Radical-Stroke Index" (<https://www.unicode.org/versions/UnihanCore2020RSIndex.pdf>) appears to be issue-free.