

ISO/IEC JTC1/SC2/**WG2 N4969**
 ISO/IEC JTC1/SC2/WG2/**IRG N2298R2**
L2/18-066R2

Universal Multiple-Octet Coded Character Set
International Organization for Standardization

Doc Type: ISO/IEC JTC1/SC2/WG2
Title: Proposed kllCore Changes & Additions
Source: Ken Lunde (Adobe) & Jaemin Chung
Status: Individual Contribution
Action: For consideration by WG2
Date: 2018-06-01 (see [L2/18-066](#) & [L2/18-066R](#))

This is a second revision of the original proposal (aka IRG N2298 & [L2/18-066](#)) that incorporates feedback from discussions during UTC #155, along with feedback from Jaemin Chung (aka [L2/18-071](#)), and now incorporates responses to [ROK's feedback](#).

As documented in a recent five-part series on the [CJK Type Blog](#), we propose the following changes be made to the [kllCore](#) property, which are based on obvious omissions (untagged KS X 1001 ideographs) or updated versions of regional standards that clearly influence the IICore repertoire, such as Japan's Jōyō Kanji and Hong Kong SAR's HKSCS-2016:

1. The latest IICore, [Version 2.2](#) (June 2004), has been stable for nearly 15 years. As a result, and for the purpose of backward compatibility, we personally feel that existing IICore ideographs—even if they were included in error— should be preserved.
2. Per the [Exploring IICore—Part 1](#) article, add the “K” source tag to the following 14 IICore ideographs, because they are the **only** ideographs—among the 4,620 in the KS X 1001 standard—that do not already have the “K” source tag:

Ideograph	Proposed kllCore	Ideograph	Proposed kllCore
塞 U+585E	AGTJH K MP	滲 U+6E17	AG K
獎 U+596C	A K P	耇 U+8009	A K P
復 U+5FA9	ATJH K MP	胄 U+80C4	AGTJ K P
慄 U+6144	ATJH K MP	詰 U+8A70	ATJH K MP
戀 U+6200	ATH K MP	諾 U+8AFE	ATJH K MP
撚 U+649A	ATJH K MP	輦 U+8F26	ATJH K P
栗 U+6817	AGTJH K MP	默 U+9ED8	AGTH K MP

According to Section 1.2 of ROK's feedback, five (5) of these 14 ideographs—the ones that are highlighted in the table above—do not require the “K” source tag to be added, because they claim that U+734E 獎, U+6EF2 滲, U+8008 耇, U+5191 胄, and U+9ED9 默 are the preferred (aka normalized) ideographs. While this may justify keeping the “K” source tag for these five ideographs that are outside the scope of KS X 1001, it does not even come close to justifying not adding the “K” source tag for the five that are within the scope of KS X 1001. Furthermore, four of the five non-KS X 1001 ideographs have a K2 (KS X 1027-1) source reference, and U+9ED9 默 does not have a kIRG_KSource property value at all. This effectively means that the number of Korean fonts that will include glyphs for these ideographs can be measured on one hand.

According to Section 3 of ROK's feedback, the “K” source tag should be removed from U+67BE 柿, and U+67FF 柿 should retain its “K” source tag. The former ideograph is within the scope of KS X 1001, and should therefore retain its “K” source tag, and the latter ideograph has a K2 (KS X 1027-1) source reference, which doesn't justify its priority over the former due to well-established industry practices in ROK.

3. Per the [Exploring IICore—Part 2](#) article, add the “J” source tag to the following two (2) IICore ideographs, because they were added to Japan’s Jōyō Kanji (常用漢字) at the end of 2010:

Ideograph	Proposed kIICore	Ideograph	Proposed kIICore
剝 U+525D	AT J HKMP	頰 U+9830	AT J HKMP

4. Per the [Exploring IICore—Part 4](#) article, add the “H” and “M” source tags to the following 10 IICore ideographs, because they are included in HKSCS-2016 and their Big Five counterparts are tagged “H” and “M” (Macao SAR follows Hong Kong SAR in this regard):

Ideograph	Proposed kIICore	Ideograph	Proposed kIICore
兑 U+5151	AG HM	脱 U+8131	AG JHM
媪 U+5AAA	CG HM	蛻 U+8715	AG HM
悦 U+60A6	AG JHM	說 U+8AAC	A JHM
愠 U+6120	CG HM	銳 U+92ED	A JHM
稅 U+7A0E	AG JHM	閱 U+95B2	A JHM

5. Per the [Exploring IICore—Part 2](#) and [Part 4](#) articles, add the following six (6) ideographs to IICore as Category “A,” because they were added to Japan’s Jōyō Kanji (常用漢字) at the end of 2010 (tagged “J”), for compatibility with Big Five (tagged “T”), or they are included in HKSCS-2016 and have Big Five counterparts (tagged “H”):

Ideograph	Proposed kIICore	Ideograph	Proposed kIICore
告 U+543F	AH	蘊 U+85F4	AH
彝 U+5F5E	AT	醞 U+9196	AH
氫 U+6C32	AH	叱 U+20B9F	AJ

6. Per [L2/18-071](#), the “K” source tag should be removed from the following 60 IICore ideographs, because they are outside the scope of the KS X 1001 standard, and no one in Korea (ROK) can provide justification as to why they were included in the first place (no documents or reports can be found):

Ideograph	Current	Proposed	Ideograph	Current	Proposed
俐 U+4FD0	AGT JHKM	AGT JHM	漳 U+6F33	AGT HKM	AGT HM
儻 U+5110	AT HKM	AT HM	漾 U+6F3E	AGT HKM	AGT HM
册 U+518C	AG K	AG	炆 U+7086	BT HKM	BT HM
胃 U+5191	AT JHK	AT JH	獎 U+734E	AT HKM	AT HM
咳 U+545F	BJ K	BJ	玕 U+7395	BT K	BT
咏 U+548F	AGT JK	AGT J	璵 U+7471	BT K	BT
屮 U+5795	BG K	BG	璦 U+7473	AT JK	AT J
垸 U+57B8	BG K	BG	璫 U+74B1	BT K	BT
堙 U+582D	BT K	BT	璜 U+74B8	BT K	BT
塹 U+587D	AT K	AT	畫 U+756B	AT HKM	AT HM
夫 U+592C	BG K	BG	祕 U+7955	AT HK	AT H
奘 U+5958	AGT JHK	AGT JH	禛 U+799B	BT K	BT
媧 U+59DE	BG K	BG	离 U+79BB	AG K	AG
婷 U+5A77	AGT HKM	AGT HM	荃 U+8343	AGT HKM	AGT HM
媿 U+5A84	BT K	BT	譏 U+8B7F	BT K	BT
媿 U+5A9C	BT K	BT	豐 U+8C50	AT HKM	AT HM
媿 U+5ABC	AT JHKM	AT JHM	適 U+9052	BG K	BG
愜 U+6103	BJ K	BJ	鈕 U+9215	AT JHKM	AT JHM

Ideograph	Current	Proposed	Ideograph	Current	Proposed
聆 U+6624	AT K	AT	鈹 U+924D	AT K	AT
哲 U+6662	BT JK	BTJ	鉑 U+9251	ATH KM	ATHM
帆 U+674B	BT K	BT	鉕 U+927C	AT K	AT
杠 U+6760	AGT K	AGT	銻 U+92D5	BT K	BT
柿 U+67FF	AGTJH KM	AGTJHM	銻 U+9315	ATH K	ATH
栞 U+681E	BJ K	BJ	銻 U+93F5	BTH KM	BTHM
桃 U+6844	BT K	BT	鑲 U+9436	BT K	BT
樞 U+6A7F	AJ K	AJ	鑊 U+943A	ATH KM	ATHM
檠 U+6AA0	AGT K	AGT	隼 U+96BC	AGT JK	AGTJ
汙 U+6C84	BGT K	BGT	靚 U+975A	BTH KM	BTHM
涼 U+6DBC	ATJH KM	ATJHM	頤 U+9824	ATJH KM	ATJHM
滲 U+6EF2	ATJH KMP	ATJHMP	默 U+9ED9	AJ K	AJ

Per the second and third paragraphs of Section 2 of this document, ROK's response may justify retaining the "K" source tag for the five highlighted ideographs, though the authors of this proposal are still dubious of their claim.

7. Also per [L2/18-071](#), and as a result of the discussions that transpired during UTC #155, the "K" source tag should be changed to "X" for the following 78 IICore ideographs that would have been left tag-less as a result of removing the "K" source tag, because they are also outside the scope of the KS X 1001 standard, and no one in Korea (ROK) can provide justification as to why they were included in the first place (no documents or reports can be found):

Ideograph	Current	Proposed	Ideograph	Current	Proposed
悒 U+3960	CK	CX	檣 U+6A72	CK	CX
秣 U+4137	CK	CX	泮 U+6CD9	CK	CX
翳 U+4A12	CK	CX	涎 U+6D8F	CK	CX
頤 U+4AB8	CK	CX	激 U+6EB5	CK	CX
倂 U+4F56	CK	CX	滢 U+6F0C	CK	CX
倂 U+4FD9	CK	CX	潤 U+6F63	CK	CX
冏 U+518F	CK	CX	焯 U+70D3	CK	CX
垌 U+578C	CK	CX	焯 U+7180	CK	CX
垌 U+579E	CK	CX	熾 U+71B2	CK	CX
埠 U+57FB	CK	CX	熾 U+7214	CK	CX
塢 U+5823	CK	CX	珧 U+73BF	CK	CX
塢 U+5890	CK	CX	琬 U+73D7	CK	CX
塢 U+58A1	CK	CX	琅 U+73E2	CK	CX
侏 U+593D	CK	CX	琬 U+73EB	CK	CX
媯 U+59B8	CK	CX	琬 U+73F5	CK	CX
媯 U+5A0D	CK	CX	瑀 U+73F8	CK	CX
媯 U+5A2B	CK	CX	璩 U+741F	CK	CX
媯 U+5A60	CK	CX	璩 U+7445	CK	CX
媯 U+5A64	CK	CX	璩 U+7446	CK	CX
媯 U+5A93	CK	CX	璩 U+744C	CK	CX
媯 U+5AB4	CK	CX	璩 U+747D	CK	CX
媯 U+5AD9	CK	CX	璩 U+74A4	CK	CX
媯 U+5AE4	CK	CX	璩 U+74D3	CK	CX

Ideograph	Current	Proposed	Ideograph	Current	Proposed
嶠 U+5D44	CK	CX	碇 U+7807	CK	CX
𣎵 U+5F54	CK	CX	袖 U+79DE	CK	CX
恣 U+5FDE	CK	CX	穊 U+7A66	CK	CX
𣎵 U+615C	CK	CX	𣎵 U+8008	CK	CX
懍 U+6193	CK	CX	𣎵 U+8740	CK	CX
𣎵 U+62AE	CK	CX	𣎵 U+8ADD	CK	CX
𣎵 U+6543	CK	CX	𣎵 U+8CF0	CK	CX
𣎵 U+65F2	CK	CX	𣎵 U+8CF1	CK	CX
𣎵 U+6621	CK	CX	𣎵 U+9250	CK	CX
𣎵 U+6679	CK	CX	𣎵 U+9277	CK	CX
𣎵 U+668B	CK	CX	𣎵 U+92E5	CK	CX
𣎵 U+66A3	CK	CX	𣎵 U+9327	CK	CX
𣎵 U+67D6	CK	CX	𣎵 U+9351	CK	CX
𣎵 U+683B	CK	CX	𣎵 U+93CB	CK	CX
𣎵 U+686F	CK	CX	𣎵 U+9442	CK	CX
𣎵 U+6A5A	CK	CX	𣎵 U+962D	CK	CX

Per the second paragraph of Section 2 of this document, ROK’s response may justify retaining the “K” source tag for the highlighted ideograph, though the authors of this proposal are still dubious of their claim.

8. The data file named “unihan-changes.txt” that is an attachment to this document provides the changes and additions to the Unihan Database as described in Sections 2 through 7 of this proposal. Of the 170 records, 26 are for adding one or two source tags, six are new records, 60 have the “K” source tag removed, and 78 have the “K” source tag replaced by the new “X” source tag.
9. Per Section 7, the Syntax field of the kIICore property as described in [UAX #38](#) should be changed to the following in order to accommodate the new “X” source tag that is used to provide a source when an existing sole source has been removed:
[ABC] ([GHJKMPT] {1, 7} | X)
10. Also per Section 7, the Description field of the kIICore property as described in [UAX #38](#) should be changed to reflect the new “X” source tag, such as by adding the following sentence: *The letter “X” is used as the source when the sole source has been removed.*

That is all.