

Universal Multiple-Octet Coded Character Set  
UCS

ISO/IEC JTC1/SC2/WG2/IRG N2079

Date: 2015-08-19

Source:	Japan
Title:	Concern about inconsistency on Source Separation
Meeting:	IRG#44
Status :	FYI
Actions required	for consideration
Distribution:	IRG
Medium :	Electronic
Page:	8
References:	Annex S of ISO/IEC 10646

On DAM2 text of ISO/IEC 10646 4th edition, it is proposed that the source reference of CJK ideographs with J1 source change to the one with JIS X0213:2004 source as below. Note that Japan voted against this proposal on the ballot of PDAM2.2 and PDAM2.3, but the PDAM was approved.

<i>Add the following entries:</i>	
J3B	JIS X 0213:2004 level-3 characters replacing J1 characters
J3C	JIS X 0213:2004 level-3 characters replacing JA characters
J4B	JIS X 0213:2004 level-4 characters replacing J1 characters
J4C	JIS X 0213:2004 level-4 characters replacing JA characters
<i>Replace all J1 and JA sources that are also included in JIS X 0213:2004 by the values defined in JIS X 0213:2004 using the prefixes described above (J3B, J3C, J4B, and J4C). This represents 85 original JA sources and 2743 original J1 sources.</i>	

*(Extracted from DAM ballot text for amendment 2)*

This change will have an impact to the unification procedure because J1 is one of the source standards that the source separation rule is applied while JIS X0213 is not. When two characters with similar shapes are encoded separately, it would be difficult to examine whether they are separated because of the unification rule or because of the source separation rule after J1 source reference changed to JIS X0213 source reference. Japan has concern about this. IRG needs to be careful when discussing unification issue. Therefore, Japan is preparing adding some note to Annex S and Annex P to make this situation clear.

However, the same situation already exists regarding T sources because TE is changed to T3 and T4 in the version published around year 2000, without any record. In the current version of the standard, T source references are described as below.

The Hanzi T sources are

T1	TCA-CNS 11643-1992 1st plane
T2	TCA-CNS 11643-1992 2nd plane
T3	TCA-CNS 11643-1992 3rd plane with some additional characters
T4	TCA-CNS 11643-1992 4th plane
T5	TCA-CNS 11643-1992 5th plane
T6	TCA-CNS 11643-1992 6th plane
T7	TCA-CNS 11643-1992 7th plane
TB	TCA-CNS 11643-2007 11th plane
TC	TCA-CNS 11643-2007 12th plane
TD	TCA-CNS 11643-2007 13th plane
TE	TCA-CNS 11643-2007 14th plane
TF	TCA-CNS 11643-2007 15th plane

Out of these T source standards, T1, T2 and TE are the standards that the source separation rule is applied, though the version of standards is revised.

#### S.1.6 Source separation rule

To preserve data integrity through multiple stages of code conversion (commonly known as "round-trip integrity"), any ideographs that are separately encoded in any one of the source standards listed below have not been unified.

G-source: GB2312-80, GB12345-90, GB7589-87\*, GB7590-87\*, GB8565-88\*, General Purpose Hanzi List for Modern Chinese Language\*

T-source: TCA-CNS 11643-1986/1st plane, TCA-CNS 11643-1986/2nd plane, TCA-CNS 11643-1986/14th plane\*

J-source: JIS X 0208-1990, JIS X 0212-1990

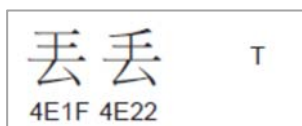
K-source: KS X 1001:2004 (previously KS C 5601-1989), KS X 1002:2001 (previously KS C 5657-1991)

NOTE 1 – A \* \* \* after the reference number of a standard indicates that some of the ideographs included in that standard are not introduced into the unified collection.

However, some ideographs encoded in two standards belonging to the same source group (e.g. GB2312-80 and GB12345-90) have been unified during the process of collecting ideographs from the source group.

The source separation rule described in S.1.6 only applies to the CJK UNIFIED IDEOGRAPHS block specified in the Basic Multilingual Plane.

Examples in S.3 (Source separation examples) are not consistent with this text. For example,



This example is the first pair in S.3. Glyph shapes are basically unifiable, but they are encoded separately because of Source separation. However, in the current code chart, u+4e22 is:



The source separation rule should not be applied to T3 source. This may cause confusion to people who reads ISO/IEC 10646.

Japan expects that TCA will also request to WG2 to add some information about this situation.

Appendix 1 Possible inconsistent J sources after Amendment 2 is published

填填	TJ
5861 586B	

[U+5861] J1-3834 → [J3B-2F58](#) [U+586B] J0-4536

擊擊	TJ
6483 64CA	

[U+6483] J0-3762 [U+64CA] J1-4123 → [J3B-7522](#)

瘦瘦	J
75E9 7626	

[U+75E9] J0-4169 [U+7626] J1-4D77 → [J3BA-7E7D](#)

蓄蓄	TJ
83D1 8458	

[U+83D1] J1-583B → [J3B-7B21](#) [U+8458] J1-5858

蔣蔣	GJ
848B 8523	

[U+848B] J0-3E55 [U+8523] J1-5936 → [J3B-7B36](#)

醬醬	J
91A4 91AC	

[U+91A4] J0-3E5F [U+91AC] J1-6273 → [J3B-7C79](#)

顛顛	J
985A 985B	

[U+985A] J1-6833 → [J3B-7E23](#) [U+985B] J0-453F

Note: Source indexes are taken from CJKSrc.txt of DAM2 text of ISO/IEC10646:2014

Appendix 2 inconsistent T-sources in Annex S.3 examples

SRC	UCS1	UCS2	UCS3	T1	T2	T3
T	04E1F	04E22		T1-4723	T3-2263	
GT	04E48	05E7A		T1-443D	T3-213D	
GTJ	04E89	0722D		T3-2265	T1-4E68	
T	04F75	05002		T1-4B77	T3-2F4B	
T	04FA3	04FB6		T3-274D	T1-4F73	
TJK	04FC1	04FE3		T2-2925	T3-2B35	
T	04FDE	0516A		T1-4F7E	T3-2B37	
T	04FF1	05036		T1-5434	T3-2F50	
T	05024	0503C		T3-2F4E	T1-542B	
T	05077	05078		T1-5922	T3-344C	
TJ	0507D	050DE		T1-5873	T3-4576	
T	0514C	05151		T1-492B	T3-2451	
TJ	0514E	05154		T3-2454	T1-4C22	
T	05156	05157		T3-2752	T1-5022	
TJ	0518A	0518C		T1-4575	T3-2179	
T	051E2	051E3		T3-212F	T3-2135	
TJ	05203	05204		T1-4443	T3-2136	
TJ	0520A	0520B		T1-457A	T3-217D	
T	05220	0522A		T3-2461	T1-4934	
T	05225	0522B		T1-4931	T3-2460	
TJ	05238	052B5		T1-4C2C	T3-2761	
T	05239	0524E		T3-275F	T1-5026	
T	0524F	05259		T3-2B43	T3-2F62	
T	0525D	05265		T1-544C	T3-2F63	
T	052FB	05300		T1-4524	T3-2156	
T	05355	05358		T3-2763	T3-2B4A	
TK	05373	0537D		T1-493B	T3-2B4B	
TJ	05377	05DFB		T1-4C3A	T3-2C44	
GT	053C1	053C2		T3-2769	T3-276A	
T	053C3	053C4		T1-5936	T3-3455	
T	05415	05442		T3-233B	T1-4948	
T	0541E	05451		T1-493F	T3-2522	
TJ	05433	05434	05449	T1-4946	T3-247B	T3-247C
T	05436	05450		T1-4951	T3-247D	
T	0543F	0544A		T3-2521	T1-494B	
T	05527	0559E		T1-5463	T3-3A40	
T	055A9	055BB		T3-3A42	T1-5E51	
T	05618	05653		T3-4633	T1-6C38	
GTJ	0568F	05694		T1-734A	T3-575B	
T	056EF	056FD		T3-2529	T3-277B	
TJ	05708	0570F		T1-594E	T3-3A43	
T	0570E	05713		T3-3A44	T1-6351	
T	05716	05717		T1-6837	T3-4637	
T	05759	05DE0		T3-645F	T2-236A	
T	05848	0588D		T2-3A69	T3-4645	
TJ	05861	0586B		T3-4058	T1-6358	
T	05897	0589E		T3-4640	T1-6C41	

SRC	UCS1	UCS2	UCS3	T1	T2	T3
GTJ	058EE	058EF		T3-2343	T1-4967	
T	058FD	05900		T1-6840	T3-6466	
T	05910	0657B		T3-4646	T2-5067	
GTJ	05932	0672C		T3-2237	T1-465B	
TJ	05968	0596C	0734E	T3-4061	T3-4648	T1-6D64
GT	05986	0599D		T3-234D	T1-4969	
T	0598D	059F8		T1-4970	T3-2C2B	
T	059CD	059D7		T1-4C6D	T3-2846	
GT	059EB	059EC		T3-2C23	T1-5478	
T	05A1B	05A2F	05A31	T1-5476	T3-304D	T3-304B
T	05A55	05AAB		T2-3453	T3-3A78	
T	05A7E	05AAE		T3-3A7C	T2-3B25	
TK	05AAA	05ABC		T3-3A79	T1-6365	
T	05AAF	05B00		T2-3A77	T3-4B5E	
T	05B0E	05B14		T3-4B5F	T2-565F	
GT	05B24	05B37		T1-7351	T3-5453	
T	05B73	05B76		T1-5E6B	T3-4072	
T	05BAB	05BAE		T3-2C32	T1-5527	
T	05BDB	05BEC		T3-4076	T1-6C4F	
T	05BDC	05BE7		T3-4073	T1-684F	
GTJ	05BDD	05BE2		T4-3C6E	T1-6854	
GTJ	05C06	05C07		T3-3059	T1-5972	
T	05C13	05C14		T4-2173	T3-223F	
T	05C19	05C1A		T3-2851	T1-4C7E	
T	05C2A	05C2B		T2-235F	T3-255F	
T	05C36	05C37		T3-5454	T1-7353	
T	05C4F	05C5B		T1-506E	T3-3543	
GT	05CE5	05D22		T3-2C43	T1-597B	
T	05DD3	05DD4		T3-5F43	T1-7B5D	
T	05E21	05E32		T2-2A31	T3-355B	
TJ	05E2F	05E36		T3-306A	T1-5A27	
T	05E76	05E77		T1-477E	T3-2863	
T	05EC4	05ECF		T1-5F23	T3-4661	
T	05F11	05F12		T3-3B42	T1-636F	
T	05F37	05F3A		T1-5A30	T3-3B47	
T	05F39	05F3E		T3-3563	T3-3B44	
TJ	05F50	05F51		T3-2140	T3-2141	
T	05F54	05F55		T2-2655	T3-2871	
T	05F59	05F5A		T1-6370	T3-4123	
T	05F5D	05F5E		T3-5764	T1-7641	
T	05F65	05F66		T1-507D	T3-2C50	
T	05FB3	05FB7		T3-4663	T1-6C61	
T	05FB4	05FB5		T3-4664	T1-6C62	
TJ	06075	060E0		T3-307B	T1-5F2D	
T	06085	060A6		T1-5550	T3-3123	
T	0609E	060AE		T3-3127	T3-3126	
T	060B3	060EA		T3-3B54	T3-3B4D	
T	06120	0614D		T3-3B62	T1-6421	

SRC	UCS1	UCS2	UCS3	T1	T2	T3
TJ	0613C	0614E		T3-466C	T1-637C	
GT	06229	0622C		T2-4A3D	T3-4675	
T	0622F	06231		T3-4C2C	T3-506D	
T	06236	06237	06238	T1-4542	T3-2165	T3-2166
T	0623B	0623E		T3-2628	T1-4D53	
T	0629B	062CB		T3-2633	T1-4D61	
TJ	0629C	062D4		T3-2634	T1-4D60	
T	06329	0635D		T2-2F40	T3-3136	
TJ	0633F	063D2	063F7	T3-313C	T1-5F46	T3-3B72
TJ	0634F	063D1		T1-555C	T3-3B73	
TJ	0635C	0641C		T3-313B	T1-6432	
T	063B2	063ED		T3-362D	T1-5F4B	
TJ	063FA	06416	06447	T4-376F	T1-6436	T3-4145
T	063FE	06435		T3-3B6F	T2-4335	
TJ	06483	064CA		T3-4C35	T1-7364	
T	0654E	06559		T3-3636	T1-5A74	
T	06553	0655A		T2-3569	T3-3634	
T	065E2	065E3		T1-514D	T3-3641	
T	06602	0663B		T1-4D7B	T3-2D22	
T	0665A	06669		T1-5B25	T3-3C26	
T	066A8	066C1		T1-6924	T3-5123	
T	067B4	067FA		T1-5165	T2-2B40	
T	067E5	067FB		T1-5167	T3-2D34	
T	067F5	06805		T1-5160	T3-2D38	
T	068B2	068C1		T2-3623	T3-3656	
T	06961	06986		T3-4179	T1-6454	
T	06982	069EA		T1-644D	T3-4C60	
T	06985	069B2		T3-4172	T3-473E	
T	0699D	06A27		T3-474E	T2-5122	
TJ	069D8	06A23		T3-4C54	T1-6D35	
T	06A2A	06A6B		T3-4C59	T1-7077	
T	06B65	06B69		T1-4A63	T3-2956	
T	06B72	06B73		T1-6459	T3-4221	
T	06B7F	06B81		T1-4E3F	T3-2958	
GTJ	06BBB	06BBC		T3-3665	T1-6026	
T	06BC0	06BC1		T1-645A	T3-4226	
T	06BCE	06BCF		T3-237A	T1-4A64	
T	06C32	06C33		T3-422A	T1-693B	
T	06C5A	06C61		T3-2423	T1-4844	
TJ	06C92	06CA1		T1-4A74	T3-2659	
TJ	06D44	06DE8		T3-2D55	T1-5B67	
T	06D89	06E09		T1-5644	T3-3725	
T	06D97	06D9A		T2-303B	T3-317B	
T	06D99	06DDA		T3-317D	T1-5B61	
T	06DE5	06E0C		T2-3655	T3-367D	
T	06DF8	06E05		T3-367C	T1-5B52	
T	06E07	06E34		T3-3724	T1-603E	
T	06E29	06EAB		T3-3C6B	T1-646B	

SRC	UCS1	UCS2	UCS3	T1	T2	T3
T	06E88	06F59		T2-3D31	T3-4C7D	
T	06E89	06F11		T1-6046	T3-4764	
T	06EDA	06EFE		T3-423E	T1-693E	
GTJK	06F5B	06FF3		T1-6D51	T3-5A55	
T	07028	0702C		T1-7842	T3-5A53	
GTJ	070BA	07232		T1-5233	T3-3D22	
GTJK	0712D	07162		T3-3D21	T2-4458	
T	07174	07185		T3-4253	T2-4B64	
GT	072B6	072C0		T3-266B	T1-4E6D	
TJ	07464	07476		T1-6964	T3-4823	
T	074F6	07501		T1-5C25	T3-4277	
T	07522	07523		T1-5C28	T3-3745	
T	076A1	076A5		T3-5176	T3-5557	
TJ	0771E	0771F		T3-3759	T1-5678	
TJK	0773E	08846		T1-5C38	T3-3E6A	
T	07814	0784F		T1-5266	T3-3761	
TJ	0797F	07984		T1-655A	T3-3D63	
T	079BF	079C3		T1-4B35	T3-2A3E	
T	07A05	07A0E		T1-6124	T3-3D68	
TJ	07A42	07A57		T3-4D57	T1-7454	
T	07BB3	07C08		T3-4D5D	T3-5570	
T	07BE1	07C12		T1-7163	T3-5571	
T	07CA4	07CB5		T3-3E23	T1-656D	
T	07D55	07D76		T1-613A	T3-3E33	
T	07DA0	07DD1		T1-6A3A	T3-486B	
T	07DD2	07DD6		T1-6A45	T3-4E22	
T	07DE3	07E01		T1-6E3F	T3-4D7E	
T	07DFC	07E15		T3-4E28	T2-5941	
T	07E48	07E66		T1-7470	T3-585A	
TJ	07FAE	07FB9		T3-4E31	T1-7867	
T	07FF6	07FFA		T3-557E	T3-5861	
T	080FC	08141		T1-5766	T3-3E45	
T	0812B	08131		T1-5C6B	T3-3845	
T	0817D	08183		T3-4422	T2-4D2D	
GT	08203	08204		T3-3E49	T2-3F3C	
TJ	0820D	0820E		T1-4F3C	T3-2A5D	
TJ	08358	0838A		T3-3363	T1-5C7D	
TJ	083D1	08458		T2-3F67	T3-4440	
T	08480	08495		T3-4439	T3-4933	
T	0848D	0853F		T2-4654	T3-525A	
T	08570	08580		T3-5255	T2-5E61	
T	085AB	085B0		T3-562C	T1-7734	
T	085F4	0860A		T3-5B25	T1-7A38	
T	0865A	0865B		T3-3861	T1-6170	
T	086FB	08715		T1-6647	T3-4446	
TJK	0885B	0885E		T1-6E6C	T3-5263	
TK	0886E	0889E		T3-3374	T1-5D35	
T	08A2E	08A7D		T4-353D	T3-445C	

SRC	UCS1	UCS2	UCS3	T1	T2	T3
T	08AAA	08AAC		T1-6B29	T3-4966	
TJ	08ACC	08AEB		T3-4E69	T1-7239	
T	08C5C	08C63		T2-3952	T3-4464	
TJ	08D70	08D71		T1-4B4C	T3-243A	
T	08EFF	08F27		T2-477B	T3-4F35	
T	08F3C	08F40		T3-5325	T2-5F6D	
T	08FBE	08FD6		T3-2730	T2-2870	
TJ	08FF8	0902C		T1-5853	T3-3F51	
T	090A2	090C9		T1-4B58	T3-2F3C	
T	090CE	090DE		T1-5367	T3-3429	
T	090F7	09109	09115	T3-3952	T1-6254	T3-4525
T	09196	0919E		T3-532F	T1-756A	
T	09203	09292		T2-412C	T3-4A33	
T	092B3	092ED		T1-6F5E	T3-4F4C	
T	09304	09332		T1-7267	T3-533A	
TK	0932C	0934A		T3-533B	T1-7570	
TJ	093AD	093AE		T3-5944	T1-7758	
T	095B1	095B2		T1-6F65	T3-4F55	
T	09751	09752		T3-2A7E	T1-4F65	
GTJ	09759	0975C		T3-4A4C	T1-727B	
T	09839	0983D		T1-7325	T3-535F	
TJ	0984F	09854		T1-776D	T3-5956	
TJ	09905	09920		T1-6B6F	T3-5364	
TJK	099B1	099C4		T1-676A	T3-4A6A	
TK	099E2	09A08		T1-732E	T3-596A	
T	09AA9	09AAB		T3-402F	T2-486F	
T	09AD8	09AD9		T1-586D	T4-362D	
TJ	09AEA	09AEE		T3-4A72	T1-7021	
T	09B2C	09B2D		T3-612E	T4-6D37	
TJ	09C1B	09C2E		T3-5D70	T3-5F2A	
T	09CEF	09CF3		T3-4567	T1-6B7A	
T	09EAA	09EAB		T3-502F	T3-502E	
T	09EBC	09EBD		T1-6B7B	T3-4B22	
T	09EC3	09EC4		T1-6327	T3-396A	
T	09ED1	09ED2		T1-6329	T3-396B	

(End of Document)