

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization

**Doc Type:** Working Group Document

**Title:** Stabilizing CJK Compatibility Ideographs through the use of the IVD

**Authors:** Ken Lunde

**Source:** The Unicode Consortium

**Status:** Liaison Contribution

**Action:** For consideration by JTC1/SC2/WG2

**Date:** 2011-10-19

## Background

CJK Compatibility Ideographs, for which there are now 1,002 characters as of Unicode Version 6.1, are subject to normalization, and are thus considered unstable because the distinctions that they are intended to convey cannot be preserved, regardless of the normalization form. When normalized, a CJK Compatibility Ideograph is reverted into its canonical equivalent, which is always a CJK Unified Ideograph. For example, the CJK Compatibility Ideographs U+FA5D (<sup>+</sup><sub>+</sub>) and U+FA5E (<sup>+</sup><sub>+), when normalized, are reverted to their shared canonical equivalent, specifically U+8279 (<sup>+</sup><sub>+</sub>). Furthermore, given the broad extent to which text services interact in today's applications and OSeS, it is not possible to guarantee that normalization will not be applied, except for completely closed environments.</sub>

The highest-frequency CJK Compatibility Ideographs are those used by Japan, and the UTC has identified 89 among the 1,002 CJK Compatibility Ideographs that fall into this category of significant real-world usage. 75 of these CJK Compatibility Ideographs correspond to JIS X 0213 characters, and 57 of those 75 JIS X 0213 characters correspond to *Jinmei-yō Kanji* (人名用漢字). Furthermore, these 89 CJK Compatibility Ideographs require 92 unique glyphs for faithful representation.

The IVD (*Ideographic Variation Database*)<sup>\*</sup> offers a solution whose representation is stable, because it is not subject to normalization. The 92 required glyphs for the 89 CJK Compatibility Ideographs were registered on 2007-12-14 as part of the "Adobe-Japan1" IVD Collection. Nearly the same set of glyphs were submitted for registration on 2011-07-05 (aka, PRI 187<sup>†</sup>) as part of the "Hanyo-Denshi" IVD Collection.

## Recommendation

We hereby request that WG2 draft a resolution to further investigate this issue, and if deemed necessary, to declare and establish a stable representation for the 1,002 CJK Compatibility Ideographs through the use of the IVD, beginning with the 89 that have been identified as being critical for Japanese use. Because a single standardized representation is needed for the 92 glyphs required for Japanese use, there are effectively three choices: 1) use the registered "Adobe-Japan1" IVSes (*Ideographic Variation Sequences*); 2) use the soon-to-be-registered "Hanyo-Denshi" IVSes; or 3) register a completely new set of IVSes, through a new IVD collection, for this specific purpose. We recommend choice #1, primarily on the grounds that there already exists a non-trivial number of fonts that implement these IVSes in their Format 14 'cmap' subtables.

The table that spans the following two pages provides the details for choices #1 and #2 as they apply to these 89 CJK Compatibility Ideographs.

\* <http://www.unicode.org/ivd/>

† <http://www.unicode.org/ivd/pri187/>

CJK Compatibility Ideograph	Canonical Equivalent	Adobe-Japan1 IVS	Adobe-Japan1 identifier	JIS X 0213	Jinmei-yō Kanji	IBM	ARIB STD-B24	Hanyo-Denshi IVS	Hanyo-Denshi identifier	PRI 187
U+F91D	U+6B04	<6B04,E0100>	CID+13392	1-86-27	JINMEI					JC8627
U+F928	U+5ECA	<5ECA,E0101>	CID+20303	1-84-14	JINMEI					JC8414
U+F929	U+6717	<6717,E0100>	CID+20305	1-85-46	JINMEI					JC8546
U+F936	U+865C	<865C,E0100>	CID+13394	1-91-47	JINMEI					JC9147
U+F970	U+6BBA	<6BBA,E0100>	CID+13344	1-86-41						JC8641
U+F9D0	U+985E	<985E,E0100>	CID+13396	1-94-04	JINMEI					JC9404
U+F9DC	U+9686	<9686,E0101>	CID+13393	1-93-61						JC9361
U+FA10	U+585A	<585A,E0101>	CID+7746	1-15-55						JC1555
U+FA10	U+585A	<585A,E0102>	CID+8422			IBM				IB1603
U+FA12	U+6674	<6674,E0100>	CID+8481			IBM				IB2015
U+FA15	U+51DE	<51DE,E0101>	CID+20307	1-87-58						JC8758
U+FA15	U+51DE	<51DE,E0104>	CID+8542			IBM				IB2339
U+FA16	U+732A	<732A,E0100>	CID+8548	1-87-79	JINMEI	IBM				JC8779
U+FA17	U+76CA	<76CA,E0101>	CID+8571			IBM				JTFA17
U+FA18	U+793C	<793C,E0101>	CID+8579			IBM				IB2536
U+FA19	U+795E	<795E,E0100>	CID+8580	1-89-28	JINMEI	IBM				JC8928
U+FA1A	U+7965	<7965,E0100>	CID+8581	1-89-29	JINMEI	IBM				JC8929
U+FA1B	U+798F	<798F,E0101>	CID+8583	1-89-33	JINMEI	IBM				JC8933
U+FA1C	U+9756	<9756,E0101>	CID+8587			IBM				IB3208
U+FA1D	U+7CBE	<7CBE,E0100>	CID+8590			IBM				JTFA1D
U+FA1E	U+7FBD	<7FBD,E0100>	CID+8599			IBM				IB2730
U+FA20	U+8612	<8612,E0101>	CID+8612			IBM				JTBAAD
U+FA20	U+8612	<8612,E0102>	CID+21073	2-87-24						JD8724
U+FA22	U+8AF8	<8AF8,E0100>	CID+8622	1-92-14	JINMEI	IBM				JC9214
U+FA25	U+9038	<9038,E0102>	CID+8633			IBM				JTFA25S
U+FA26	U+90FD	<90FD,E0100>	CID+8636	1-92-74	JINMEI	IBM				JC9274
U+FA2A	U+98EF	<98EF,E0100>	CID+8699			IBM				JTFA2A
U+FA2B	U+98FC	<98FC,E0101>	CID+8700			IBM				JTFA2B
U+FA2C	U+9928	<9928,E0101>	CID+8702			IBM				IB0457
U+FA2D	U+9DB4	<9DB4,E0100>	CID+8715			IBM				IB1173
U+FA30	U+4FAE	<4FAE,E0101>	CID+13382	1-14-24	JINMEI					JC1424
U+FA31	U+50E7	<50E7,E0101>	CID+13360	1-14-41	JINMEI					JC1441
U+FA32	U+514D	<514D,E0101>	CID+13389	1-14-48						JC1448
U+FA33	U+52C9	<52C9,E0100>	CID+13385	1-14-67	JINMEI					JC1467
U+FA34	U+52E4	<52E4,E0101>	CID+13338	1-14-72	JINMEI					JC1472
U+FA35	U+5351	<5351,E0100>	CID+13378	1-14-78	JINMEI					JC1478
U+FA36	U+559D	<559D,E0101>	CID+7651	1-15-12						JC1512
U+FA37	U+5606	<5606,E0100>	CID+13366	1-15-15	JINMEI					JC1515
U+FA38	U+5668	<5668,E0101>	CID+13333	1-15-22	JINMEI					JC1522
U+FA39	U+5840	<5840,E0101>	CID+13384	1-15-58						JC1558
U+FA3A	U+58A8	<58A8,E0100>	CID+13387	1-15-62	JINMEI					JC1562
U+FA3B	U+5C64	<5C64,E0101>	CID+13361	1-47-65	JINMEI					JC4765
U+FA3C	U+5C6E	<5C6E,E0100>	CID+16837	1-47-66						JC4766
U+FA3D	U+6094	<6094,E0100>	CID+13326	1-84-48	JINMEI					JC8448
U+FA3E	U+6168	<6168,E0101>	CID+13328	1-84-60						JC8460
U+FA3F	U+618E	<618E,E0100>	CID+13363	1-84-62	JINMEI					JC8462

CJK Compatibility Ideograph	Canonical Equivalent	Adobe-Japan1 IVS	Adobe-Japan1 identifier	JIS X 0213	Jinmei-yō Kanji	IBM	ARIB STD-B24	Hanyo-Denshi IVS	Hanyo-Denshi identifier	PRI 187
U+FA40	U+61F2	<61F2,E0101>	CID+13369	1-84-65	JINMEI					JC8465
U+FA41	U+654F	<654F,E0100>	CID+13381	1-85-08	JINMEI					JC8508
U+FA42	U+65E2	<65E2,E0100>	CID+13334	1-85-11						JC8511
U+FA43	U+6691	<6691,E0101>	CID+13352	1-85-35	JINMEI					JC8535
U+FA44	U+6885	<6885,E0100>	CID+13375	1-85-69	JINMEI					JC8569
U+FA45	U+6D77	<6D77,E0100>	CID+13327	1-86-73	JINMEI					JC8673
U+FA46	U+6E1A	<6E1A,E0100>	CID+7700	1-86-87	JINMEI					JC8687
U+FA47	U+6F22	<6F22,E0101>	CID+13332	1-87-05	JINMEI					JC8705
U+FA48	U+716E	<716E,E0100>	CID+13347	1-87-53	JINMEI					JC8753
U+FA49	U+722B	<722B,E0101>	CID+15398	2-80-09						JD8009
U+FA4A	U+7422	<7422,E0100>	CID+7732	1-88-05	JINMEI					JC8805
U+FA4B	U+7891	<7891,E0100>	CID+13379	1-89-07	JINMEI					JC8907
U+FA4C	U+793E	<793E,E0101>	CID+13348	1-89-19	JINMEI					JC8919
U+FA4D	U+7949	<7949,E0101>	CID+13345	1-89-20	JINMEI					JC8920
U+FA4E	U+7948	<7948,E0100>	CID+13335	1-89-23	JINMEI					JC8923
U+FA4F	U+7950	<7950,E0100>	CID+13391	1-89-24	JINMEI					JC8924
U+FA50	U+7956	<7956,E0101>	CID+13359	1-89-25	JINMEI					JC8925
U+FA51	U+795D	<795D,E0100>	CID+13351	1-89-27	JINMEI					JC8927
U+FA52	U+798D	<798D,E0100>	CID+13325	1-89-31	JINMEI					JC8931
U+FA53	U+798E	<798E,E0101>	CID+13371	1-89-32	JINMEI					JC8932
U+FA54	U+7A40	<7A40,E0100>	CID+13343	1-89-45	JINMEI					JC8945
U+FA55	U+7A81	<7A81,E0101>	CID+13373	1-89-49	JINMEI					JC8949
U+FA56	U+7BC0	<7BC0,E0101>	CID+13358	1-89-68	JINMEI					JC8968
U+FA57	U+7DF4	<7DF4,E0100>	CID+13399	1-90-14	JINMEI					JC9014
U+FA58	U+7E09	<7E09,E0101>	CID+18366	2-84-48						JD8448
U+FA59	U+7E41	<7E41,E0101>	CID+13376	1-90-19	JINMEI					JC9019
U+FA5A	U+7F72	<7F72,E0100>	CID+13353	1-90-26	JINMEI					JC9026
U+FA5B	U+8005	<8005,E0101>	CID+13349	1-90-36	JINMEI					JC9036
U+FA5C	U+81ED	<81ED,E0101>	CID+13350	1-90-56	JINMEI					JC9056
U+FA5D	U+8279	<8279,E0101>	CID+14199	2-85-84						JD8584
U+FA5E	U+8279	<8279,E0102>	CID+14198	2-85-85						JD8585
U+FA5F	U+8457	<8457,E0101>	CID+13367	1-91-07	JINMEI					JC9107
U+FA60	U+8910	<8910,E0101>	CID+13331	1-91-79						JC9179
U+FA61	U+8996	<8996,E0101>	CID+13346	1-91-89	JINMEI					JC9189
U+FA62	U+8B01	<8B01,E0100>	CID+13321	1-92-15	JINMEI					JC9215
U+FA63	U+8B39	<8B39,E0100>	CID+13339	1-92-16	JINMEI					JC9216
U+FA64	U+8CD3	<8CD3,E0100>	CID+13380	1-92-24	JINMEI					JC9224
U+FA65	U+8D08	<8D08,E0101>	CID+13364	1-92-29	JINMEI					JC9229
U+FA66	U+8FB6	<8FB6,E0101>	CID+15403	2-89-73						JD8973
U+FA67	U+9038	<9038,E0101>	CID+13320	1-92-57	JINMEI					JC9257
U+FA68	U+96E3	<96E3,E0100>	CID+13374	1-93-67	JINMEI					JC9367
U+FA69	U+97FF	<97FF,E0101>	CID+13337	1-93-86	JINMEI					JC9386
U+FA6A	U+983B	<983B,E0100>	CID+7788	1-93-91						JC9391
U+FA6B	U+6075	<6075,E0101>	CID+13740				ARIB			JTB16C
U+FA6C	U+242EE	<242EE,E0101>	CID+14281				ARIB			JTB538
U+FA6D	U+8218	<8218,E0101>	CID+13695				ARIB	<8218,E0103>	IB1067	