ISO/IEC JTC1/SC2/WG2
Universal Multiple-Octet Coded Character Set (UCS) – ISO/IEC 10646
Secretariat: USA

TITLE:	Report of DPRK compatibility characters ad hoc meeting
SOURCE:	DPRK compatibility characters ad hoc group
STATUS:	Expert Contribution
DISTRIBUTION:	Members of ISO/IEC JTC1/SC2/WG2
DATE:	2002-12-11
REFERENCE:	**WG2 N2375, N2493, N2540, N2541, N2566**

During the WG2 meeting #43, DPRK ad hoc group consisting of experts from relevant national bodies was organized to address the issues, primarily raised by the US asking for removal of DPRK compatibility characters from FA70 to FAE9 as part of FPDAM ballot comments.

This is a summary report from the DPRK compatibility characters ad hoc group.

1. Errors in FPDAM code table containing 122 DPRK compatibility characters

The group reviewed the document from the contributing editor and confirmed that the corrected code table for DPRK compatibility characters in WG2N2540 is exactly what is intended and should have been used for FPDAM.   Those errors in FPDAM code table were introduced during editing process and it doesn't mean the original input from DPRK was unstable.

2. Categorization of Comments

The group reviewed 23 characters according to WG2N2566 and 2 characters in US comment (T.8) and found two of them are just a mere mistake in WG2N2566.

As a result of our discussion on the rest of 23 characters, the group addressed four categories of issues regarding these characters.

Type-1: Considering the nature of compatibility characters being used solely for round-trip mapping, we should make use of assigned code positions for compatibility characters of one national standard source for those compatibility characters from another national standard as long

as the same unified counter part is shared and there's no confusion for the intended shape.

Type-2: Even if the same shape of a proposed compatibility character happens to be found in existing CJK unified repertoire and its extensions, in principle we should not use the code position in the unified repertoire, or otherwise the normalization information between the intended variant and the unified counter part will be lost.

Type-3: If we find the unifiable shape of a proposed compatibility character in CJK Ext-C candidates and if the Ext-C candidate is considered a new character in terms of unification procedure adopted by the IRG, then the same character should not be encoded as a compatibility character.

Type-4:   If a proposed compatibility character is considered a new character in terms of unification procedure adopted by the IRG and it does not appear in the Ext-C candidates, then the character should be proposed as an Ext-C candidate and should not be encoded as a compatibility character.

3. Recommendation from the Ad hoc Group

Based on the discussion according to four categories described above, the group recommends removing 15 characters from the original proposal for encoding 122 DPRK compatibility characters and encode the rest of 107 DPRK compatibility characters in 10646-1 Amendment 2.

The list of 15 characters for removal is as follows:

Code positions in WG2N2540	KP code	Category
FA7C	KP1-3A92	Type-1
FA88	KP1-42C8	Type-1
FA8C	KP1-43C1	Type-1
FA8D	KP1-43D2	Type-1
FA93	KP1-45D5	Type-1
FA9C	KP1-4B5D	Type-1
FAA5	KP1-54B7	Type-1
FAB7	KP1-639A	Type-1
FABC	KP1-6725	Type-1
FAD1	KP1-8054	Type-1
FAD2	KP1-80AD	Type-1
FAE1	KP1-57D0	Type-1
FAE6	KP1-78AC	Type-1
FA90	KP1-441D	Type-4
FA9F	KP1-510B	Type-4

More detail information about characters discussed by the ad hoc group is attached in this document.

Attachment

Total # of characters reviewed: 25

WG2N2566 errors: 2

  FA96, FA7E

C1 candidates: 8

1. Existing C1candidates should be unified: 6

   The following six characters were claimed to be unified to Ext-C candidates. Ext-C is now under extensive review and the group considered that the corresponding Ext-C candidate characters are suspicious to be remained in the final Ext-C set. The group hence recommends to keep these six characters in DPRK compatibility ideographs.

   (There were no Type-3 characters and therefore all 6 characters should remain in the compatibility character proposal.)

   FABB, FA9B, FA86, FAB3, FAAD, FA8A

2. New character proposal (Type-4): 2

   The following two characters were claimed to be rather unified ideographs than compatibility ideographs. The group recognized that there two characters are more appropriate as candidates for unified ideographs and recommends to remove from DPRK compatibility ideographs.

   FA90, FA9F

Ext-B character: 1

The character FA7D was claimed to be unified with U+20F96. The group considered that, per the unification rule for CJK unified ideogrpahs, suggested unification of FA7D and 20F96 is inappropriate. The group recommends to keep it in DPRK compatibility.

FA7D should not be mapped to U+20F96 (Type-2)

Map to plane 2 compatibility characters (Type-1): 13

The following 13 characters were claimed to be same as those in (pre-existing) compatibility ideographs. The group recognized that each of the following characters is appropriate to be

mapped to its corresponding compatibility ideograph in plane 2 and recommends to remove from DPRK compatibility ideographs.

FA7C, FA88, FA8C, FA8D, FA93, FA9C, FAA5, FAB7, FABC, FAD1, FAD2, FAE1, FAE6

Compatibility character as in the original proposal: 1

The following character was claimed to be considered as a candidate for a future extension work of unified ideographs since it cannot be unified with any of existing unified ideographs. The group considered that the following character is to be unified with 233D5 per the unification rule, if it is proposed as a unified ideograph candidate. The group hence recommends to keep the following character in DPRK compatibility ideographs.

FADF

Character by character analyses follow:

WG2N2566 errors: 2

FA96 U+6452 = U+2F8BA

Resolution: WG2N2566 indicates that U+2F8BA is a possible candidate for mapping from KP1-46F9, but they don't share the same unified counter part.

52	摒	KP1-46F9	摒	KP1-4670	U+06452

拼
2F8BA (->U+62FC)

FA7E U+585A => U+585C

Resolution: WG2N2566 indicates the unified counter part of FA7E should be U+585C, but apparently this is a false alarm.

21	塚	KP1-3BEE	塚	KP0-EEA9

088/090 585A	塚	塚	塚	塚
	8-2E32	1-6356	0-444D	0-7540
	8-1418	1-6754	0-3645	0-8532

088/092 585C	塚	塚		塚
	5-365A	3-4056		2-2A67
	5-2258	3-3254		2-1071

Existing C1 candidates: 6

FABB EXT. C2, p 423, No. 16482:

Resolution: according to Annex S, this C1 candidate should be unified with U+7DF4

練	KP1-671B	練	KP0-D8AA	U+07DF4, U+0F996

16482	zjw01422		練
17741			
0929.291			
糸	TD-5357		
8 | 1			

FA9B EXT. C2, p.216, No. 10760:

Resolution: the C1 candidate should be unified to U+6756, and add this example to the IRG's standing document on unification examples..

杖	KP1-4ABD	杖	KP0-E6EA	U+06756

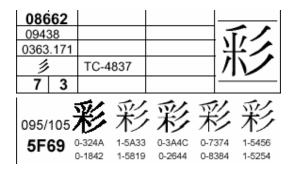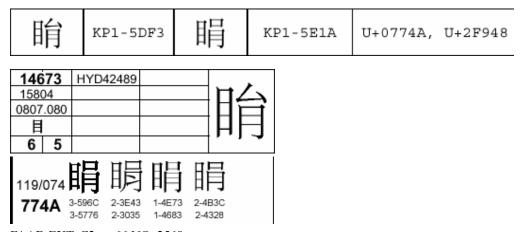| 10760			杖
11645			
0517.011			
木	TC-3346		
4 | 1			

FA86 EXT. C2, p. 223 NO. 8662

Resolution: the C1 candidate should be unified to U+5F69, and add this example to the IRG's standing document on unification examples.

彩	KP1-41FB	彩	KP0-EFBA	U+05F69

08662			彩
09438			
0363.171			
彡	TC-4837		
7	3		

095/105	彩 彩 彩 彩 彩		
5F69	0-324A 1-5A33 0-3A4C 0-7374 1-5456		
	0-1842 1-5819 0-2644 0-8384 1-5254		

FAB3 EXT. C2, p. 377 NO. 14673

Resolution: according to Annex S, this C1 candidate should be unified with U+774A

| 睅 | KP1-5DF3 | 睅 | KP1-5E1A | U+0774A, U+2F948 |

14673	HYD42489		睅
15804			
0807.080			
目			
6	5		

119/074	睅 睅 睅 睅		
774A	3-596C 2-3E43 1-4E73 2-4B3C		
	3-5776 2-3035 1-4683 2-4328		

FAAD EXT. C2, p. 66 NO. 2568:

Resolution: the pair are already unified. IRG should follow up to make sure this one is to be removed from C1 candidate.

| 画 | KP1-5ABB | 画 | KP1-5ABA | U+0753B |

02568			画
02789			
0135.161			
凵	TC-2F24		
6	1		

FA8A EXT. C2: pick one of NO.9007,9025 or 9026

Resolution: the C1 candidates should have unified counter parts, and KP1-437E should not proposed as an Ext-C candidate.

慎	KP1-437E	慎	KP1-43C0	U+0614E, U+2F8A8	

097/078 614E	慎 0-4977 0-4187	慎 1-637C 1-6792	慎 0-3F35 0-3121	慎 1-556E 1-5378

09007 09789 0396.071 心 9 | 1	TD-3A6B		慎
09025 09808 0396.071 心 9 | 3	TD-266C		慎
09026 09809 0396.071 心 9 | 3	TD-2675		慎

Ext-B character: 1

FA7D U+5668 = U+20F96

Resolution: This is Type-2. KP1-3AA3 should not be mapped to U+20F96.

17	器	KP1-3AA3	器	KP0-D2BB	U+05668, U+0FA38

086/104 5668	器 0-4977 0-3887	器 1-7042 1-8034	器 0-346F 0-2079	器 0-506F 0-4879	器 1-5021 1-4801

20F96 06623 06623 0209.151 □ 12	T4-4932		器 器 器 器

Map to plane 2 compatibility characters (Type-1): 13

FAA5 U+7070 = U+2F835

78	灰	KP1-54B7	灰	KP0-F5E8

灰
2F835

FA7C U+5606 = U+2F84C

16	嘆	KP1-3A92	嘆	KP0-F0A6	U+05606, U+0FA37, U+2F84C

嘆
2F84C

FA88 U+6081 = U+2F8A0

34	悁	KP1-42C8	悁	KP1-4302	U+06081, U+2F8A0

悁
2F8A0

FA8D U+614C = U+2F8A7

39	慌	KP1-43D2	慌	KP0-F6E1	U+0614C, U+2F8A7

慌
2F8A7

FA93 U+6350 = U+2F8B7

48	捐	KP1-45D5	捐	KP0-F8E4	U+06350, U+2F8B7

2F8B7

FA9C U+6852 = U+2F8E1

61	栞	KP1-4B5D	栞	KP1-4B07	U+06852, U+2F8E1

2F8E1

FAE1 U+24735 = U+2F926

156	犴	KP1-57D0	犴	KP1-579D	U+24735, U+2F926

24735 30283 30283 0707.010 21335.010 犬 4	G_KX T5-245A KP1-579D		2F926

FAB7 U+7AEE = U+2F95F

103	屏	KP1-639A	屏	KP1-6395	U+07AEE, U+2F95F

2F95F

FABC U+7D63 = U+2F96C

108	絣	KP1-6725	絣	KP1-66C0	U+07D63, U+2F96C

2F96C

FAE6 U+27CA8 = U+2F9D3:

159	豻	KP1-78AC	豻	KP1-789B	U+27CA8, U+2F9D3

27CA8 52865 52865 1200.110 63909.010 豸 4	G_KX T4-354E KP1-789B	豻 豻 豻 豻	豻 2F9D3

FAD1 U+92D7 = U+2F9E8

135	銗	KP1-8054	銷	KP1-8089	U+092D7, U+2F9E8

146/215 92D7	銷 銷 銷 銷 3-7639 2-5537 1-6441 1-7355 3-8625 2-5323 1-6833 1-8353	銷 2F9E8

FAD2 U+927C = U+2F9EA

136	鉼	KP1-80AD	鉼	KP1-8041	U+0927C, U+2F9EA

146/124 927C	鉼 鉼 鉼 鉼 3-7632 1-6B59 1-6375 1-6179 3-8618 1-7557 1-6785 1-6589	鉼 2F9EA

FA8C

38	慈	KP1-43C1	慈	KP0-E6B2	U+06148, U+2F8A6

097/072	慈	慈	慈	慈	慈		慈
6148	0-3448	1-6375	0-3B7C	0-6D31	1-556B		
	0-2040	1-6785	0-2792	0-7717	1-5375		2F8A6

New character proposal (Type-4): 2

FA9F != U+6C9B

Resolution: This is a non-unifiable pair in terms of unification rule adopted by the IRG. KP1-510B should be submitted to the IRG as a new character.

67	沛	KP1-510B	沛	KP0-F2A5	U+06C9B

108/155	沛	沛	沛	沛	沛
6C9B	0-4566	1-4A6C	0-5D6F	0-7829	1-5B76
	0-3770	1-4276	0-6179	0-8809	1-5986

水 4

FA90

Resolution: This is a non-unifiable pair in terms of unification rule adopted by the IRG. KP1-441D should be submitted to the IRG as a new character.

42	憮	KP1-441D	憡	KP0-CFA7	U+0614A

097/074	慊	慊	慊	慊
614A	0-633B	2-426D	0-5844	0-4C43
	0-6727	2-3477	0-5636	0-4435

Compatibility character as in the original proposal: 1

FADF != U+233D5

WG2N2566 indicates that KP1-4B26 may be a candidate as a new character, but the IRG experts confirmed that they should be unified. The IRG is asked to add this example to the standing document on unification examples..

155	枕	KP1-4B26	枕	KP1-4ACF	U+233D5

233D5	G_KX
21639	T5-264C
21639	J4-2E44
0513.070	
21165.010	
木	
4	KP1-4ACF

橔橔

699D 6A27

T