

ISO/IEC JTC1/SC2/WG2

Universal Multi-Octet Coded Character Set

(UCS)

ISO/IEC JTC1/SC2/WG2

N2048

1999.8.6

TITLE: Proposal to supplement the Arabic Coded character set with the special script and character for Uighur, Kazakh and kirghiz languages

SOURCE: CHINA

STATUS: NATIONAL BODY POSITION

ACTION: CONSIDERATION BY WG2

China's Xinjiang Uighur Autonomous Regional Peoples' Government attaches great importance to the appointed work of information processing and interchange of such information in the Uighur, Kazakh and Kirghiz and Arabic character sets using multi-Octet international standards. Since the 90's much support and help have been given to our work by the experts from various countries at the ISO/IEC JTC1/SC2/WG2 conferences. Consequently, great advancement has been made such that the majority of the characters script and graphic characters of the Uighur, Kazakh and Kirghiz languages which are used in information processing and interchange such information are now included in this standard. Therefore on behalf the Xinjiang Uighur Autonomous Regional Peoples' Government and her multi-ethnic peoples, we express our heartfelt gratitude and thanks to all of you. Now We are proposing to have a supplementary scheme for Uighur, Kazakh and Kirghiz scripts which were inadvertently excluded from this standard in 1995. Regarding the proposed supplementary scheme, we have convened many symposia since the conference in Finland in June 1995. In this symposia we have met with the relevant experts both in spoken and written languages as well as in information processing on computers and related department heads and have had serious and extended discussions about the proposal. We have finally decided on the following supplementary scheme which we hope will be given serious deliberation and approval in this conference.

PROPOSED SUPPLEMENTARY SCHEME

1. In the scheme submitted at the Conference on Oct, 8, 1996 there were 18 vowel letters with Hamza in the Uighur language. We now propose to supplement two more

vowel letters with Hamza "هـ، وْ" which are also vowel letters. With the inclusion of these two vowel letters with Hamza there will be 20 vowel letters with Hamza in the standard. Only then can we ensure that information processing and interchange in Uighur language can be perfected. If these 2 vowel letters with Hamza are not included it is impossible to display all the necessary characters used in information processing and interchange. Now we suggested that they should be placed in FBD1 and FBD2 in the table 114.

هـ (FBD1) ARABIC LIGATURE UIGHUR YEH WITH HAMZA ABOVE WITH E MEDIAL FORM.

وْ (FBD2) ARABIC LIGATURE UIGHUR YIH WITH HAMZA ABOVE WITH ALEF MAKSURA MEDIAL FORM.

2. In the scheme submitted in Oct 1996, We have proposed to supplement 12 different phonemes corresponding to the 12 graphic characters in the Kazakh language. Presently we have abandoned this original scheme and will not be submitting it for approval.

3. The 2 graphic form "ي، يْ" which have long been used in the Kirghiz language were proposed to be included in the standard in 1996. This is because these 2 graphic forms are different in appearance from their 2 corresponding characters in table 114, FE89, FE8A. Also these "ي، يْ" graphic forms cannot be substituted by the FE89, FE8A corresponding characters. Therefore these 2 graphic forms need to be used independently and should be included in the standard. Now we suggested that they should be placed in FBCF and FBD0 in the table 114.

ي (FBCF) ARABIC LIGATURE KIRGHIZ YEH WITH HAMZA ABOVE ISOLATED FORM.

يْ (FBD0) ARABIC LIGATURE KIRGHIZ YEH WITH HAMZA ABOVE FINAL FORM.

Based on past research, we propose to supplement these 6 ligatures of the long vowels "ا، آ، آٲ، آٳ، آٴ، آٵ" in the Kirghiz language. Reason: There are 30 letters representing 36 phonemes in the Kirghiz language. These 6 long vowels have not been included in the alphabetic system. Since 1990, these 6 long vowels have not been listed in any of the proposed schemes for the character set in the Kirghiz language. The reason being that attention was focused on the nominal forms and the various appearance for the letters of this language, to the neglect of the existence of these 6 long vowels' (ligatures) in the language. As a result there were no graphic forms produced for these 6 long vowels. Therefore, the 6 long vowels could not be accurately inputted, displayed and printed correctly during information processing and information interchanges. Facts have proven the necessity of supplementing these 6 ligatures to facilitate proper and

effective information processing and interchange in the Kirghiz language. Thus we would especially request the inclusion of these 6 ligatures in the supplementary scheme for the Kirghiz language in 1996 and have it coded in FD40--FD45 in table 117.

- 11 (FD40) ARABIC LIGATURE KIRGHIZ ALEF WITH ALEF ISOLATED FORM.
- oo (FD41) ARABIC LIGATURE KIRGHIZ AE WITH AE ISOLATED FORM.
- ââ (FD42) ARABIC LIGATURE KIRGHIZ YU WITH YU ISOLATED FORM.
- ûû (FD43) ARABIC LIGATURE KIRGHIZ U WITH U ISOLATED FORM.
- œœ (FD44) ARABIC LIGATURE KIRGHIZ OE WITH OE ISOLATED FORM.
- 99 (FD45) ARABIC LIGATURE KIRGHIZ WAW WITH WAW ISOLATED FORM.

In conclusion, we would like to thank all of you for the opportunity to present these supplementary schemes for Uighur and kirghiz languages. We hope that they would meet with your approval and henceforth be included in the standard.

Note: "Proposal in 1996" referred in this paper is numbered WG2 N2027.

表 114 行 FB: 阿拉伯文字变形显现形式-A
TABLE 114 - Row FB: ARABIC PRESENTATION FORMS-A

	FB8	FB9	FBA	FBB	FBC	FBD	FBE	FBF
0	چ 129	ک 144	ن 160	ل 176		عی 202	و 224	ئو 240
1	چ 129	ک 145	ن 161	ل 177		ئبد 203	و 225	ئو 241
2	د 130	گ 146	ن 162			ئد 204	و 226	ئو 242
3	د 131	گ 147	ن 163			ئد 205	و 227	ئو 243
4	د 132	گ 148	ه 164			ئد 206	پ 228	ئو 244
5	د 133	گ 149	ه 165			ئد 207	پ 229	ئو 245
6	ذ 134	گب 150	ه 166			ئد 208	پ 230	ئی 246
7	ذ 135	گب 151	ه 167			ف 209	پ 231	ئی 247
8	ذ 136	گب 152	ر 168			ف 210	ر 232	ئی 248
9	ذ 137	گب 153	ر 169			ف 211	ر 233	ئی 249
A	ر 138	نگ 154	ه 170			ف 212	ئا 234	ئی 250
B	ر 139	نگ 155	ه 171			ف 213	ئا 235	ئی 251
C	ر 140	نگ 156	ه 172			ف 214	ئا 236	ئی 252
D	ر 141	نگ 157	ه 173			ف 215	ئا 237	ئی 253
E	ک 142	ن 158	ل 174			و 222	ئو 238	ی 254
F	ک 143	ن 159	ل 175		عی 202	و 223	ئو 239	ی 255

G=00

P=00

表 117 行 FD: 阿拉伯文字变形呈现形式—A
TABLE 117 — Row FD: ARABIC PRESENTATION FORMS-A

	FD0	FD1	FD2	FD3	FD4	FD5	FD6	FD7
0	جي	ضر	خي	ش	١١	ن	ح	ك
	000	016	032	048		080	096	112
1	جى	طى	صى	س	هه	ن	ج	ط
	001	017	033	049		081	097	113
2	جى	طى	صى	ش	وؤ	ن	ج	ط
	002	018	034	050		082	098	114
3	جى	مى	ضى	ط	وؤ	ن	ج	ط
	003	019	035	051		083	099	115
4	جى	مى	ضى	سج	وؤ	ن	ج	ط
	004	020	036	052		084	100	116
5	صى	مى	ش	س	وؤ	ن	ج	ط
	005	021	037	053		085	101	117
6	جى	مى	ش	س		ن	ج	ط
	006	022	038	054		086	102	118
7	ضى	مى	ش	س		ن	ج	ط
	007	023	039	055		087	103	119
8	ضى	مى	ش	س		ن	ج	ط
	008	024	040	056		088	104	120
9	ش	شى	شر	ش		ن	ج	ط
	009	025	041	057		089	105	121
A	ش	شى	سر	ط		ن	ج	ط
	010	026	042	058		090	106	122
B	ش	جى	ضر	ظ		ن	ج	ط
	011	027	043	059		091	107	123
C	ش	جى	ضر	أ		ن	ج	ط
	012	028	044	060		092	108	124
D	شر	جى	ش	أ		ن	ج	ط
	013	029	045	061		093	109	125
E	سر	جى	ش	و		ن	ج	ط
	014	030	046	062		094	110	126
F	ضر	جى	ش	و		ن	ج	ط
	015	031	047	063		095	111	127

G = 00

P = 00