Doc Type	
Title:	**Proposal to add a Named UCS Sequence Identifier UYGHUR LETTERS**
Source:	China
Action:	**Review at WG2 Meeting**
Distribution:	**SC2/WG2**

The Unicode Technical Committee received input from Uyghur experts indicating the cultural importance of a text element referred to as vowel letters:

	Nominal form	Alphabet letters name	sequence characters	sequence characters code	Isolated form and represented code	Initial form and represented code	Medial form and represented code	Final form and represented code
1	ئا	Uyghur letter A	ئ+ا	0626+0627	ئا FBEA			ئا FBEB
					ا FE8D			ا FE8E
2	ئە	Uyghur letter AE	ئ+ە	0626+06D5	ئە FBEC			ئە FBED
					ە FEE9			ە FEEA
3	ئې	Uyghur letter E	ئ+ې	0626+06D0	ئې FBF6	ئﯦ FBF8	ﯦ FBD1	ئې FBF7
					ﯦ FBE4	ﯦ FBE6	ﯧ FBE7	ﯦ FBE5
4	ئى	Uyghur letter I	ئ+ى	0626+06D9	ئى FBF9	ئﯨ FBFB	ﯨ FBD2	ئى FBFA
					ى FEEF	ﯨ FBE8	ﯩ FBE9	ى FEF0
5	ئو	Uyghur letter O	ئ+و	0626+0648	ئو FBEE			ئو FBEF
					و FEED			و FEEE
6	ئۇ	Uyghur letter U	ئ+ۇ	0626+06C7	ئۇ FBF0			ئۇ FBF0
					ۇ FBF0			ۇ FBF0
7	ئۆ	Uyghur letter OE	ئ+ۆ	0626+06C6	ئۆ FBF2			ئۆ FBF3
					ۆ FBD9			ۆ FBDA
8	ئۈ	Uyghur letter OU	ئ+ۈ	0626+06C8	ئۈ FBF4			ئۈ FBF5
					ۈ FBDB			ۈ FBDC
9	نک	Chaghatay letter NG	ن+ک	0646+06A9	نک FBD1	نک FBD2	نک FBD0	نک FBCF

This text element is perceived culturally to be a single character and is listed and named in "alphabet" lists. This text element has a representation in the UCS as a character sequence: <0626,0627>. And every vowel letters linker mode.

On that basis, the Unicode Consortium proposes addition of the following named UCS sequence identifier in Amendment 8 of ISO/IEC 10646:

Uyghur letter A : Nominal form ئا , sequence character <ئ+ا>, sequence code <0626,0627> .

      Isolated form ئا , sequence character <ئ+ا>, sequence code <0626,0627>, represented code < FBEA >

      Isolated form ا, sequence character <ا >, sequence code <0627>, represented code < FE8D >

      final form ئا , sequence character <ئ+ا>, sequence code <0626,0627>, represented code < FBEB >

      final form ا, sequence character <ا >, sequence code <0627>, represented code < FE8E >

Uyghur letter AE : Nominal form ئە , sequence character <ئ+ە>, sequence code <0626,06D5> .

      Isolated form ئە , sequence character <ئ+ە>, sequence code <0626,06D5>, represented code < FBEC >

      Isolated form ە, sequence character <ە >, sequence code <06D5>, represented code < FEE9 >

      final form ئە , sequence character <ئ+ە>, sequence code <0626,06D5>,   represented code < FBED >

      final form ە, sequence character <ە >, sequence code <06D5>, represented code < FEEA >

Uyghur letter E : Nominal form ئې , sequence character <ئ+ې>, sequence code <0626,06D0> .

      Isolated form ئې, sequence character <ئ+ې>, sequence code<0626,06D0>, represented code < FBF6 >

      Isolated form ې, sequence character <ې >, sequence code<06D0>,   represented code < FBE4 >

      Initial form ئې , sequence character <ئ+ې>, sequence code <0626,06D0>, represented code < FBF8 >

      Initial form ې , sequence character <ې >, sequence code<06D0>,   represented code < FBE6 >

      Medial form ئې , sequence character <ئ+ې>, sequence code <0626,06D0>, represented code < FBD1 >

      Medial form ې , sequence character <ې >, sequence code<06D0>, represented code < FBE7 >

      final form ئې , sequence character <ئ+ې>, sequence code <0626,06D0>, represented code < FBF7 >

      final form ې , sequence character <ې >, sequence code<06D0>, represented code < FBE5 >

Uyghur letter I : Nominal form ئى , sequence character <ئ+ى>, sequence code <0626, 06D9> .

      Isolated form ئى, sequence character <ئ+ى>, sequence code<0626, 06D9>, represented code < FBF9 >

      Isolated form ى, sequence character <ى >, sequence code <06D9>,represented code < FEEF >

      Initial form ئى , sequence character <ئ+ى>, sequence code <0626, 06D9>, represented code < FBFB >

      Initial form ى , sequence character <ى >, sequence code < 06D9>, represented code < FBE8 >

      Medial form ئى , sequence character <ئ+ى>, sequence code<0626, 06D9>, represented code <FBD2 >

      Medial form ى , sequence character <ى >, sequence code <06D9>,presented code < FBE9 >

      final form ئى , sequence character <ئ+ى>, sequence code <0626, 06D9>, represented code < FBFA >

      final form ى , sequence character <ى >, sequence code<06D9>, represented code < FEF0 >

Uyghur letter O : Nominal form ئو , sequence character <ئ+و>, sequence code <0626,0648> .

      Isolated form ئو , sequence character <ئ+و>, sequence code <0626,0648>, represented code < FBEE >

Isolated form و, sequence character <و >, sequence code <0648>, represented code < FEED >

final form ئو , sequence character <و+ئ>, sequence code <0626,0648>, represented code < FBEF >

final form و, sequence character <و >, sequence code <0648>, represented code < FEEE >

Uyghur letter U : Nominal form ئۇ , sequence character <ۇ+ئ>, sequence code <0626,06C7> .

Isolated form ئۇ , sequence character <ۇ+ئ>, sequence code <0626,06C7>, represented code < FBF0 >

Isolated form ۇ, sequence character <ۇ >, sequence code <06C7>, represented code < FBD7 >

final form ئۇ , sequence character <ۇ+ئ>, sequence code <0626,06C7>, represented code < FBF1 >

final form ۇ, sequence character <ۇ >, sequence code <06C7>, represented code < FBD8 >

Uyghur letter OE : Nominal form ئۆ , sequence character <ۆ+ئ>, sequence code <0626,06C6> .

Isolated form ئۆ , sequence character <ۆ+ئ>, sequence code <0626,06C6>, represented code < FBF2 >

Isolated form ۆ, sequence character <ۆ >, sequence code <06C6>, represented code < FBD9 >

final form ئۆ , sequence character <ۆ+ئ>, sequence code <0626,06C6>, represented code < FBF3 >

final form ۆ, sequence character <ۆ >, sequence code <06C6>, represented code < FBDA >

Uyghur letter OU : Nominal form ئۈ , sequence character <ۈ+ئ>, sequence code <0626,06C8> .

Isolated form ئۈ , sequence character <ۈ+ئ>, sequence code <0626,06C8>, represented code < FBF4 >

Isolated form ۈ, sequence character <ۈ >, sequence code <06C8>, represented code < FBDB >

final form ئۈ , sequence character <ۈ+ئ>, sequence code <0626,06C8>, represented code < FBF5 >

final form ۈ, sequence character <ۈ >, sequence code <06C8>, represented code < FBDC >

Chaghatay letter NG : Nominal form نک , sequence character <ک+ن>, sequence code <0646,06A9> .

Isolated form نک, sequence character <ک+ن>, sequence code <0646,06A9>,represented code <FBD1 >

Initial form نک , sequence character <ک+ن>, sequence code <0646,06A9>, represented code < FBD2 >

Medial form نک , sequence character <ک+ن>, sequence code <0646,06A9>, represented code < FBD0 >

final form نک , sequence character <ک+ن>, sequence code <0646,06A9>, represented code < FBCF >

Reference 1：

Table 1-2: Special vowel Scripts to Be Supplemented in   ISO/IEC 10646 /Unicode for Uyghur languages

Language	Final form	Medial form	Initial form	Isolated form	Nominal form	No
Uyghur	ا ئا FE8E FBEB			ا ئا FE8D FBEA	ئا	1
Uyghur	ه ئە FEEA FBED			ئە ه FBEC FEE9	ئە	2
Uyghur	ي ئې FBE5 FBF7	ئـ ئـ FBE7 FBD1	ئـ ئـ FBE6 FBF8	ي ئې FBE4 FBF6	ئې	3
Uyghur	ى ئى FEF0 FBFA	ـ ئـ FBE9 FBD2	ئـ ئـ FBE8 FBFB	ى ئى FEEF FBF9	ئى	4
Uyghur	و ئو FEEE FBEF			و ئو FEED FBEE	ئو	5
Uyghur	ۇ ئۇ FBD8 FBF1			ۇ ئۇ FBD7 FBF0	ئۇ	6
Uyghur	ۋ ئۆ FBDA FBF3			ۆ ئۆ FBD9 FBF2	ئۆ	7

Uyghur	ئو ؤ FBDC FBF5			ؤ ۇؤ FBDB FBF4	ۇؤ	8

	FB5	FB6	FB7	FB8	FB9	FBA	FBB	FBC	FBD	FBE	FBF
0	أ FB50	ڠ FB60	ٯ FB70	ڇ FB80	ک FB90	ݨ FBA0	ﮰ FBB0	▨	▨	و FBE0	ئۇ FBF0
1	أ FB51	ݡ FB61	ٯ FB71	ڇ FB81	ک FB91	ݩ FBA1	ﮱ FBB1	▨	▨	و FBE1	ئۇ FBF1
2	ٻ FB52	ٿ FB62	ڃ FB72	ڍ FB82	گ FB92	ۀ FBA2	▨	▨	ۇ FBE2	ئۆ FBF2
3	ٻ FB53	ٿ FB63	ڃ FB73	ڍ FB83	گ FB93	ۀ FBA3	لأ FBD3	ۇ FBE3	ئۆ FBF3
4	ٻ FB54	ٿ FB64	ڃ FB74	ڌ FB84	گ FB94	ة FBA4	لث FBD4	ى FBE4	ئۈ FBF4
5	ٻ FB55	ٿ FB65	ڃ FB75	ڌ FB85	ک FB95	ة FBA5	ݣ FBD5	ى FBE5	ئۈ FBF5
6	پ FB56	ٹ FB66	ڄ FB76	ڎ FB86	گ FB96	ه FBA6	ک FBD6	ﯦ FBE6	ئى FBF6
7	پ FB57	ٹ FB67	ڄ FB77	ڎ FB87	گ FB97	ﮧ FBA7	ۇ FBD7	ﯧ FBE7	ئى FBF7
8	پ FB58	ٹ FB68	ڄ FB78	ڈ FB88	ڰ FB98	ﮨ FBA8	ۇ FBD8	ﯨ FBE8	ئﮧ FBF8
9	چ FB59	ٺ FB69	ڄ FB79	ڈ FB89	گ FB99	ﮩ FBA9	ۈ FBD9	ﯩ FBE9	ئى FBF9
A	ﭚ FB5A	ڦ FB6A	ڇ FB7A	ڑ FB8A	گ FB9A	ه FBAA	ﯚ FBDA	ﯪ FBEA	ﯺ FBFA
B	ﭛ FB5B	ڦ FB6B	چ FB7B	ڑ FB8B	گ FB9B	ﮫ FBAB	ﯛ FBDB	ﯫ FBEB	ﯻ FBFB
C	ﭜ FB5C	ڦ FB6C	چ FB7C	ڒ FB8C	گ FB9C	ﮬ FBAC	ﯜ FBDC	ﯬ FBEC	ى FBFC
D	ﭝ FB5D	ڦ FB6D	چ FB7D	ڔ FB8D	گ FB9D	ﮭ FBAD	ﯝ FBDD	ﯭ FBED	ى FBFD
E	ٽ FB5E	ڧ FB6E	چ FB7E	ک FB8E	ں FB9E	ﮮ FBAE	ﯞ FBDE	ﯮ FBEE	ﯾ FBFE
F	ٽ FB5F	ڨ FB6F	چ FB7F	ک FB8F	ں FB9F	ﮯ FBAF	ﯟ FBDF	ﯯ FBEF	ﯿ FBFF

Reference 2： ISO/IEC JTC1/SC2/WG2 58 meeting   8 June 2011

ISO/IEC JTC1/SC2/WG2 N4113
Ad hoc report on Uyghur
SOURCE: Tero Aalto, CSC – IT Center for Science
DATE: 8 June 2011
LOCATION: Finnish Standards Association SFS, Helsinki, Finland
PARTICIPANTS:
Tero Aalto (Finland)
Chen Zhuang (China)
Mike Ksar (independent)
Lu Qin (HKSAR)
Mikami Yoshiki (Japan)
Wushouer Silamu (China)
V. S. Umamaheswaran (Canada)
Yasen Yiming (China)
The ad hoc group met and discussed issues relating to three documents:

the original proposal from China ("Preliminary Proposal for Encoding Special Scripts and Characters in UCS for Uyghur, Kazakh and Kirgiz" N3819)

comments on this proposal (N3889)

a new proposal from China ("Proposal to Encode Special Scripts and Characters in UCS for Uyghur language" N3919).

The discussion in the ad hoc group was based on the following document:

100078 N4067 – Proposal to Encode Special Scripts and Characters in UCS for Uyghur language

The eight characters proposed in N4067 can presently be produced as combinations of two existing characters. According to the proposal, these characters should, in addition, be encoded separately.
An argument supporting the proposal was the fact that the characters can have different
numbers of shapes in different languages. However, this is not within the scope of UCS. Instead, different glyph variants are handled by fonts. According to the principle, anything that can be represented with a sequence will not be encoded as a separate character.

The ad hoc group arrived at the following conclusions:

100078 There is to be a named UCS Sequence Identifier (USI) for each of the proposed

sequences.
o   A USI is also to be determined for the character proposed in N4066 –
PROPOSING TO SUPPLEMENT WITH THE SCRIPT AND CHARACTER OF Chaghatay
Language.

100078 The contributor is requested to produce and submit another contribution defining the nine

sequences for consideration in a future meeting.