

2011-11-07

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document

Title: Proposal for encoding the Uyghur script in the SMP of the UCS

Source: Global Information Infrastructure Lab, Nagaoka University of Technology, Japan

Author: Omarjan Osman

Action: For consideration by UTC and WG2

Distribution: ISO/IEC JTC 1/SC 2/WG 2, ISO/IEC JTC 1/SC 2 and Liaison Organizations

Date: 2011-11-07

Contents

Abstract.....	2
1. Introduction.....	2
1.1 The origin and the family line of Uyghur language.....	2
1.2 The origin and the system of a Uyghur characters.....	2
1.3 The classification of characters.....	3
2. Proposed character set.....	3
2.1 Classification of glyph.....	4
2.2 Ligature.....	5
3. Problem.....	5
4. Solution technique.....	6
5. Design of Uyghur Character Code Table.....	7
6. Design of Uyghur Glyph and Ligature.....	7
7. Proposed Character Code, Glyph and Ligature.....	8
7.1 Proposed Uyghur Character Code Table.....	8
7.2 Proposed Uyghur Glyph Code Table.....	9
7.3 Proposed Uyghur Ligature Code Table.....	11
8. Uyghur alphabet Vowels and Consonant.....	12
9. Uyghur alphabet and glyph types.....	12
10. Bibliography.....	13
11. Appendix: Proposal summary form.....	14

Abstract

Uyghur script which is based on Aramaic alphabet is composed of phonetic characters. The script is written horizontally, from right to left and vertically, written from top to bottom. Character structure is composed of horizontally and vertically vowels, horizontally and vertically consonants, diacritics, punctuation symbols, digits, signature symbols, glyphs and ligatures. Beforehand, the Uyghur character codes are not yet included in International standard ISO / IEC 10646 and Unicode. In the current study, the authors propose a design of Uyghur alphabets, character code, glyph and ligatures.

1. Introduction

1.1 The origin and the family line of Uyghur language

Uyghur people live Around Tangri 'heaven' Mountain range in Central Asia and Altai Mountain range. The Uyghur language belongs to the Altaic language family. Uyghur (In Uyghur character: **ئۇيغۇر**) is the name of the People. The Tangri Mountains was a part of the Uyghur country. It is called the Uyghur kingdom. The language of the Uyghur kingdom is called the old Churuk language. The Uyghur language consists of eight vowels: a --- æ, e --- i, o --- ö, u --- ü; and 24 consonants: b, p, t, dʒ, ʃ, x, d, r, z, ʒ, s, ʃ, ɣ, f, q, k, g, ŋ, l, m, n, h, v, y.

1.2 The origin and the system of a Uyghur characters

Character used by Uyghur people have changed over time, from Orkhon script in the ancient times to currently used Arabic script. Uyghur script itself became completely extinct today. Uyghur script has been existent from 7th century AD to 19th century AD in the Central Asia. The Uyghur script is derived from the Aramaic script. itself the forerunner of the Mongolian , Manchu (Sibe), Buryat and Kalmyk (Oirat) Script. In the 7th century AD, Turfan Uyghur kingdom used the Uyghur characters as the national character. Around 12th century AD, the Uyghur alphabet was later used by the Kingdom of Mongolia. Influenced by the Islamic religion in the 15th century AD, Turfan Uyghurs, and Mongols adopted the Arabic alphabet into Turfan Uyghur script.

1.3 The classification of characters

Today this Uyghur script is not used in written language. The Uyghur character, is written, horizontal writing from the right to left and vertical writing from top to bottom. Uyghur script basic set has 8 horizontal vowels letters and 8 vertical vowels letters, 24 horizontal consonant letters and 24 vertical consonant letters, diacritics, punctuation symbols, digits, signature symbols. As shown in the following definitions and abbreviations of these scripts; The vowel character is expressed "V", Horizontal Vowels characters "HV", Vertical Vowels characters "VV", Consonants "C", Horizontal Consonants characters "HC", Vertical Consonants characters "VC", Diacritic Marks "D", Punctuation Marks "P", Digits "I", Signature Symbols "J". (See table1)

Table1. Uyghur characters classification

Acronym	Name
V	Horizontal Vowels (HV)
	Vertical Vowels (VV)
C	Horizontal Consonants (HC)
	Vertical Consonants (VC)
D	Horizontal Diacritics Marks (HD)
	Vertical Diacritics Marks (VD)
P	Horizontal Punctuation Marks (HP)
	Vertical Punctuation Marks (VP)
I	Horizontal Digits (HD)

	Vertical Digits (VD)
J	Horizontal Signature Symbol (HJ)
	Vertical Signature Symbol (VJ)

2. Proposed character set

The Uygur character basic set has 8 vowels, 24 consonants letters, diacritics, punctuation symbols, digits, and signature symbols. If these signs are given, the syntactic structure of a Uygur character can be written. The following describes these character sets.

Vowels: These signs are given as 8 horizontal vowels “HV” and vertical vowels “VV”, which follows the shape of vowel letters.

Uygur characters have 8 horizontal vowels “HV” and vertical vowels “VV”:

HV (X000 ... X007), VV (X090 ... X097)

Consonants: These signs are given as 24 horizontal “HC” and vertical consonants “VC”, which follows the shape of consonant letters.

Uygur characters have 24 horizontal and vertical consonants:

HC (X008 ... X01F), VC (X098 ... X0AF)

Diacritics: These signs are given as “D” and follow the shape of diacritics.

Horizontal and vertical Uygur characters have 74 Diacritic Marks:

HD (X020, X044) VD (X0B0, X0D4)

Punctuation symbols: These signs are given as “P” and follow the shape of punctuation symbols.

Horizontal and vertical Uygur characters have 32 punctuation symbols:

HP (X050 ... X05F), VP (X0E0 ... X0EF)

Digits: These signs are given as “I” and follow the shape of Digits.

Horizontal and vertical Uygur characters have 40 Digits:

HI (X060 ... X079), VI (X0F0 ... X109)

Signature symbols: Three Signature symbols “J” follow the shape of Signature symbols.

Horizontal and vertical Uygur characters have 30 Signature symbols:

HJ (X080 ... XX08E), VJ (X110 ... XX11E)

2.1 Classification of glyph

Uygur character is a descendant of the Aramaic alphabet, and inherited Semitic system characters. The same characters change shape depending on the position of the word. All vowel and consonant letters of Uygur characters have four display forms of initial (Left-joining glyph form), medial (Dual-joining glyph form), final (Right-joining glyph form) and isolated form (Nominal glyph form) [1].

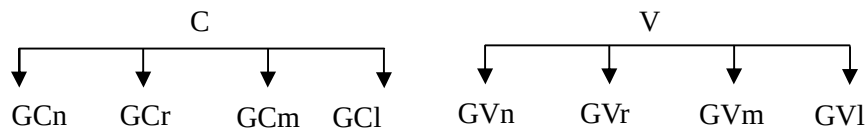


Figure1. Implementation of the Uygur Glyph

Uygur character writing is similar to Arabic. Basic principles of international standards for character codes, which have established a “Separation of the glyph code [2], are different from a code table, were designed glyphs that correspond to each display form (Glyph, abstract figure in which type is shown) [3].

2.2 Proposed Uyghur Horizontally and Vertically Glyph Table

Uyghur vowel glyphs: VG

Each vowel glyph has the following form and the following description is a record of the order of Latin characters from left to right. (See table6)

Uyghur Horizontally vowel glyphs: HVG (0000-003D)

Uyghur Vertically vowel glyphs: VVG (0100-003D)

Uyghur consonants glyphs: CG (0140 ... 01FF)

Each consonants glyph has the following form.

Uyghur Horizontally consonants glyphs: HCG (003E-00CC)

Uyghur Vertically consonants glyphs: VCG (013E-01CC)

Uyghur vowel ligature: VL

Uyghur Horizontally ligature: HL (00CF-00E9)

Uyghur Vertically ligature: VL (01CF-01E9)

3. Reference Materials

Ancient Uyghur script has been in use over thousand years since 8th century AD to 18th century AD [4]. As writing styles and letter forms have changed over time, we need to focus our scope to specific time zone and to choose specific reference material. Authors choose “Kutadgu Bilig” (means “Wisdom of Royal Glory”) written around 10th to 11th century. It was written in both ancient Uyghur script and in Arabic script. A part of the page, a horizontal and a vertical version, is shown in Fig 2. In addition to Kutadgu Bilig, two Buddhist documents [5][6] and one secular documents are also referred as a source of letter design, especially for punctuations, digits and signature symbols.



Figure 7. Uyghur Characters Vertically written from top to bottom (Kutadgu bilik 191page)

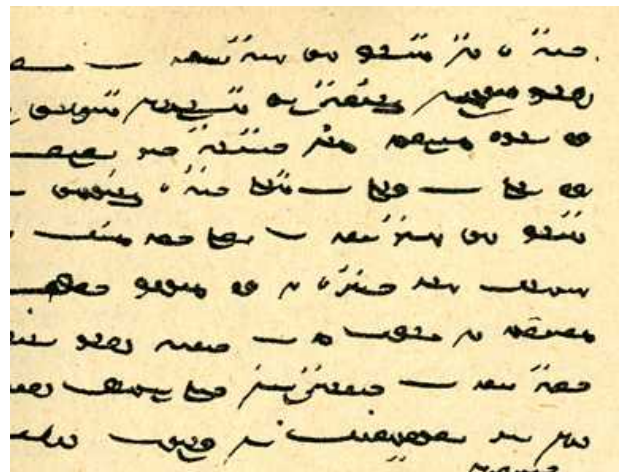


Figure 7.B Uyghur Characters Horizontally written from right to left (Kutadgu bilik 88page)

Figure2. Explanation of writing of Uyghur character

4. Proposed Uyghur Character Horizontally Code-Table

Table2. Proposed Uyghur Character Horizontally Code-Table

X000	Uyghur									X08F
	X00	X01	X02	X03	X04	X05	X06	X07	X08	
0										
1										
2										
3										
4										
5										
6										
7										
8										
9										
A										
B										
C										
D										
E										
F										

5. Proposed Uyghur Character Vertically Code-Table

Table3. Proposed Uyghur Character Vertically Code-Table

X000	Uyghur									X08F
	X09	X0A	X0B	X0C	X0D	X0E	X0F	X10	X11	
0										
	X090	X0A0	X0B0	X0C0	X0D0	X0E0	X0F0	X100	X110	
1										
	X091	X0A1	X0B1	X0C1	X0D1	X0E1	X0F1	X101	X111	
2										
	X092	X0A2	X0B2	X0C2	X0D2	X0E2	X0F2	X102	X112	
3										
	X093	X0A3	X0B3	X0C3	X0D3	X0E3	X0F3	X103	X113	
4										
	X094	X0A4	X0B4	X0C4	X0D4	X0E4	X0F4	X104	X114	
5										
	X095	X0A5	X0B5	X0C5	X0D5	X0E5	X0F5	X105	X115	
6										
	X096	X0A6	X0B6	X0C6	X0D6	X0E6	X0F6	X106	X116	
7										
	X097	X0A7	X0B7	X0C7	X0D7	X0E7	X0F7	X107	X117	
8										
	X098	X0A8	X0B8	X0C8	X0D8	X0E8	X0F8	X108	X118	
9										
	X099	X0A9	X0B9	X0C9	X0D9	X0E9	X0F9	X109	X119	
A										
	X09A	X0AA	X0BA	X0CA	X0DA	X0EA	X0FA	X10A	X11A	
B										
	X09B	X0AB	X0BB	X0CB	X0DB	X0EB	X0FB	X10B	X11B	
C										
	X09C	X0AC	X0BC	X0CC	X0DC	X0EC	X0FC	X10C	X11C	
D										
	X09D	X0AD	X0BD	X0CD	X0DD	X0ED	X0FD	X10D	X11D	
E										
	X09E	X0AE	X0BE	X0CE	X0DE	X0EE	X0FE	X10E	X11E	
F										
	X09F	X0AF	X0BF	X0CF	X0DF	X0EF	X0FF	X10F	X11F	

6. Proposed Uyghur Horizontally Glyph -Table

Table4. Proposed Uyghur Horizontally Glyph -Table

0000	Uyghur										00AF
	000	001	002	003	004	005	006	007	008	009	00A
0	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹	۰۰A0
1	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۰۰A1
2	۲۰	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶	۲۷	۲۸	۲۹	۰۰A2
3	۳۰	۳۱	۳۲	۳۳	۳۴	۳۵	۳۶	۳۷	۳۸	۳۹	۰۰A3
4	۴۰	۴۱	۴۲	۴۳	۴۴	۴۵	۴۶	۴۷	۴۸	۴۹	۰۰A4
5	۵۰	۵۱	۵۲	۵۳	۵۴	۵۵	۵۶	۵۷	۵۸	۵۹	۰۰A5
6	۶۰	۶۱	۶۲	۶۳	۶۴	۶۵	۶۶	۶۷	۶۸	۶۹	۰۰A6
7	۷۰	۷۱	۷۲	۷۳	۷۴	۷۵	۷۶	۷۷	۷۸	۷۹	۰۰A7
8	۸۰	۸۱	۸۲	۸۳	۸۴	۸۵	۸۶	۸۷	۸۸	۸۹	۰۰A8
9	۹۰	۹۱	۹۲	۹۳	۹۴	۹۵	۹۶	۹۷	۹۸	۹۹	۰۰A9
A	۰۰A0	۰۰A1	۰۰A2	۰۰A3	۰۰A4	۰۰A5	۰۰A6	۰۰A7	۰۰A8	۰۰A9	۰۰AA
B	۰۰B0	۰۰B1	۰۰B2	۰۰B3	۰۰B4	۰۰B5	۰۰B6	۰۰B7	۰۰B8	۰۰B9	۰۰AB
C	۰۰C0	۰۰C1	۰۰C2	۰۰C3	۰۰C4	۰۰C5	۰۰C6	۰۰C7	۰۰C8	۰۰C9	۰۰AC
D	۰۰D0	۰۰D1	۰۰D2	۰۰D3	۰۰D4	۰۰D5	۰۰D6	۰۰D7	۰۰D8	۰۰D9	۰۰AD
E	۰۰E0	۰۰E1	۰۰E2	۰۰E3	۰۰E4	۰۰E5	۰۰E6	۰۰E7	۰۰E8	۰۰E9	۰۰AE
F	۰۰F0	۰۰F1	۰۰F2	۰۰F3	۰۰F4	۰۰F5	۰۰F6	۰۰F7	۰۰F8	۰۰F9	۰۰AF

Table4. Uygur Horizontally Glyph Code Table

00B0	Uygur					01E9
	00B	00C	00D	00E	00F	
0						
	00B0	00C0	00D0	00E0	00F0	
1						
	00B1	00C1	00D1	00E1	00F1	
2						
	00B2	00C2	00D2	00E2	00F2	
3						
	00B3	00C3	00D3	00E3	00F3	
4						
	00B4	00C4	00D4	00E4	00F4	
5						
	00B5	00C5	00D5	00E5	00F5	
6						
	00B6	00C6	00D6	00E6	00F6	
7						
	00B7	00C7	00D7	00E7	00F7	
8						
	00B8	00C8	00D8	00E8	00F8	
9						
	00B9	00C9	00D9	00E9	00F9	
A						
	00BA	00CA	00DA	00EA	00FA	
B						
	00BB	00CB	00DB	00EB	00FB	
C						
	00BC	00CC	00DC	00EC	00FC	
D						
	00BD	00CD	00DD	00ED	00FD	
E						
	00BE	00CE	00DE	00EE	00FE	
F						
	00BF	00CF	00DF	00EF	00FF	

7. Proposed Uyghur Vertically Glyph -Table

Table5. Proposed Uyghur Vertically Glyph -Table

0100	Uyghur											01AF
	010	011	012	013	014	015	016	017	018	019	01A	
0	𐰀	𐰁	𐰂	𐰃	𐰄	𐰅	𐰆	𐰇	𐰈	𐰉	𐰊	
	0100	0110	0120	0130	0140	0150	0160	0170	0180	0190	01A0	
1	𐰋	𐰌	𐰍	𐰎	𐰏	𐰐	𐰑	𐰒	𐰓	𐰔	𐰕	
	0101	0111	0121	0131	0141	0151	0161	0171	0181	0191	01A1	
2	𐰖	𐰗	𐰘	𐰙	𐰚	𐰛	𐰜	𐰝	𐰞	𐰟	𐰠	
	0102	0112	0122	0132	0142	0152	0162	0172	0182	0192	01A2	
3	𐰡	𐰢	𐰣	𐰤	𐰥	𐰦	𐰧	𐰨	𐰩	𐰪	𐰫	
	0103	0113	0123	0133	0143	0153	0163	0173	0183	0193	01A3	
4	𐰬	𐰭	𐰮	𐰯	𐰰	𐰱	𐰲	𐰳	𐰴	𐰵	𐰶	
	0104	0114	0124	0134	0144	0154	0164	0174	0184	0194	01A4	
5	𐰷	𐰸	𐰹	𐰺	𐰻	𐰼	𐰽	𐰾	𐰿	𐱀	𐱁	
	0105	0115	0125	0135	0145	0155	0165	0175	0185	0195	01A5	
6	𐱂	𐱃	𐱄	𐱅	𐱆	𐱇	𐱈	𐱉	𐱊	𐱋	𐱌	
	0106	0116	0126	0136	0146	0156	0166	0176	0186	0196	01A6	
7	𐱍	𐱎	𐱏	𐱐	𐱑	𐱒	𐱓	𐱔	𐱕	𐱖	𐱗	
	0107	0117	0127	0137	0147	0157	0167	0177	0187	0197	01A7	
8	𐱘	𐱙	𐱚	𐱛	𐱜	𐱝	𐱞	𐱟	𐱠	𐱡	𐱢	
	0108	0118	0128	0138	0148	0158	0168	0178	0188	0198	01A8	
9	𐱣	𐱤	𐱥	𐱦	𐱧	𐱨	𐱩	𐱪	𐱫	𐱬	𐱭	
	0109	0119	0129	0139	0149	0159	0169	0179	0189	0199	01A9	
A	𐱮	𐱯	𐱰	𐱱	𐱲	𐱳	𐱴	𐱵	𐱶	𐱷	𐱸	
	010A	011A	012A	013A	014A	015A	016A	017A	018A	019A	01AA	
B	𐱹	𐱺	𐱻	𐱼	𐱽	𐱾	𐱿	𐲀	𐲁	𐲂	𐲃	
	010B	011B	012B	013B	014B	015B	016B	017B	018B	019B	01AB	
C	𐲄	𐲅	𐲆	𐲇	𐲈	𐲉	𐲊	𐲋	𐲌	𐲍	𐲎	
	010C	011C	012C	013C	014C	015C	016C	017C	018C	019C	01AC	
D	𐲏	𐲐	𐲑	𐲒	𐲓	𐲔	𐲕	𐲖	𐲗	𐲘	𐲙	
	010D	011D	012D	013D	014D	015D	016D	017D	018D	019D	01AD	
E	𐲚	𐲛	𐲜	𐲝	𐲞	𐲟	𐲠	𐲡	𐲢	𐲣	𐲤	
	010E	011E	012E	013E	014E	015E	016E	017E	018E	019E	01AE	
F	𐲥	𐲦	𐲧	𐲨	𐲩	𐲪	𐲫	𐲬	𐲭	𐲮	𐲯	
	010F	011F	012F	013F	014F	015F	016F	017F	018F	019F	01AF	

Table5. Uygur Vertically Glyph Code Table					
01B0	Uygur				01FF
	01B	01C	01D	01E	01F
0					
	01B0	01C0	01D0	01E0	01F0
1					
	01B1	01C1	01D1	01E1	01F1
2					
	01B2	01C2	01D2	01E2	01F2
3					
	01B3	01C3	01D3	01E3	01F3
4					
	01B4	01C4	01D4	01E4	01F4
5					
	01B5	01C5	01D5	01E5	01F5
6					
	01B6	01C6	01D6	01E6	01F6
7					
	01B7	01C7	01D7	01E7	01F7
8					
	01B8	01C8	01D8	01E8	01F8
9					
	01B9	01C9	01D9	01E9	01F9
A					
	01BA	01CA	01DA	01EA	01FA
B					
	01BB	01CB	01DB	01EB	01FB
C					
	01BC	01CC	01DC	01EC	01FC
D					
	01BD	01CD	01DD	01ED	01FD
E					
	01BE	01CE	01DE	01EE	01FE
F					
	01BF	01CF	01DF	01EF	01FF

8. Uygur alphabet Vowels and Consonants

The Uygur character set are used in the later stage of the script history, around the 7th century AD to 19th century AD, consists of 32 letters, 8 vowel letters plus 24 consonant letters. As shown in table 6 there are 8 horizontally vowels letter and 8 vertically vowel letter of “/a/, /æ/, /e/, /i/, /o/, /ø/, /u/, /y/”. are 24 horizontally consonants letter and 24 vertically consonants letter of “/b/, /p/, /t/, /dʒ/, /tʃ/, /x/, /d/, /r/, /z/, /ʒ/, /s/, /ʃ/, /ɣ/, /f/, /q/, /k/, /g/, /ŋ/, /l/, /m/, /n/, /h/, /v/, /j/”.

Table6. Uygur alphabet

No	Vertically Character	Horizontally Character	Sound	Name	No	Vertically Character	Horizontally Character	Sound	Name
1	ا	آ	/a/	A	17	ز	ز	/z/	ZE
2	آ	آ	/æ/	AH	18	چ	چ	/ʒ/	ZHEE
3	ئ	ئ	/e/	E	19	س	س	/s/	SI
4	ى	ى	/i/	I	20	ش	ش	/ʃ/	SHI
5	و	و	/o/	O	21	غ	غ	/ɣ/	GHE
6	ۇ	ۇ	/ø/	OV	22	ف	ف	/f/	FI
7	ۈ	ۈ	/u/	U	23	ق	ق	/q/	KHI
8	ۉ	ۉ	/y/	UV	24	ك	ك	/k/	KE
9	ب	ب	/b/	BE	25	گ	گ	/g/	GE
10	پ	پ	/p/	PE	26	ڭ	ڭ	/ŋ/	NGI
11	ت	ت	/t/	TE	27	ل	ل	/l/	LE
12	ج	ج	/dʒ/	ZHE	28	م	م	/m/	ME
13	چ	چ	/tʃ/	CHE	29	ن	ن	/n/	NE
14	خ	خ	/x/	HE	30	ھ	ھ	/h/	HHE
15	د	د	/d/	DE	31	ۋ	ۋ	/v/	VE
16	ر	ر	/r/	RE	32	ي	ي	/j/	YE

9. Uygur alphabet and glyph types

Table7. Alphabet and glyph

No	Turfan Name	Vertically				Horizontally				Sound value
		Xn	Xr	Xm	Xl	Xn	Xr	Xm	Xl	
1	A	ا	ا	ا	ا	ا	ا	ا	ا	/a/
2	AH	آ	آ	آ	آ	آ	آ	آ	آ	/æ/
3	E	ئ	ئ	ئ	ئ	ئ	ئ	ئ	ئ	/e/
4	I	ى	ى	ى	ى	ى	ى	ى	ى	/i/
5	O	و	و	و	و	و	و	و	و	/o/
6	OV	ۇ	ۇ	ۇ	ۇ	ۇ	ۇ	ۇ	ۇ	/ø/

7	U	ᱠᱨ	ᱠᱩ	ᱠᱪ	ᱠᱫ	ᱠᱬ	ᱠᱭ	ᱠᱮ	ᱠᱯ	/u/
8	UV	ᱡᱨ	ᱡᱩ	ᱡᱪ	ᱡᱫ	ᱡᱬ	ᱡᱭ	ᱡᱮ	ᱡᱯ	/y/
9	BE	ᱢᱨ	ᱢᱩ	ᱢᱪ	ᱢᱫ	ᱢᱬ	ᱢᱭ	ᱢᱮ	ᱢᱯ	/b/
10	PE	ᱣᱨ	ᱣᱩ	ᱣᱪ	ᱣᱫ	ᱣᱬ	ᱣᱭ	ᱣᱮ	ᱣᱯ	/p/
11	TE	ᱤᱨ	ᱤᱩ	ᱤᱪ	ᱤᱫ	ᱤᱬ	ᱤᱭ	ᱤᱮ	ᱤᱯ	/t/
12	ZHE	ᱥᱨ	ᱥᱩ	ᱥᱪ	ᱥᱫ	ᱥᱬ	ᱥᱭ	ᱥᱮ	ᱥᱯ	/ᄀ/
13	CHE	ᱦᱨ	ᱦᱩ	ᱦᱪ	ᱦᱫ	ᱦᱬ	ᱦᱭ	ᱦᱮ	ᱦᱯ	/ᄁ/
14	HE	ᱧᱨ	ᱧᱩ	ᱧᱪ	ᱧᱫ	ᱧᱬ	ᱧᱭ	ᱧᱮ	ᱧᱯ	/x/
15	DE	ᱨᱨ	ᱨᱩ	ᱨᱪ	ᱨᱫ	ᱨᱬ	ᱨᱭ	ᱨᱮ	ᱨᱯ	/d/
16	RE	ᱩᱨ	ᱩᱩ	ᱩᱪ	ᱩᱫ	ᱩᱬ	ᱩᱭ	ᱩᱮ	ᱩᱯ	/r/
17	ZE	ᱪᱨ	ᱪᱩ	ᱪᱪ	ᱪᱫ	ᱪᱬ	ᱪᱭ	ᱪᱮ	ᱪᱯ	/z/
18	ZHEE	ᱫᱨ	ᱫᱩ	ᱫᱪ	ᱫᱫ	ᱫᱬ	ᱫᱭ	ᱫᱮ	ᱫᱯ	/3 /
19	SI	ᱬᱨ	ᱬᱩ	ᱬᱪ	ᱬᱫ	ᱬᱬ	ᱬᱭ	ᱬᱮ	ᱬᱯ	/s/
20	SHI	ᱭᱨ	ᱭᱩ	ᱭᱪ	ᱭᱫ	ᱭᱬ	ᱭᱭ	ᱭᱮ	ᱭᱯ	/ʃ/
21	GHE	ᱮᱨ	ᱮᱩ	ᱮᱪ	ᱮᱫ	ᱮᱬ	ᱮᱭ	ᱮᱮ	ᱮᱯ	/ᵐ /
22	FI	ᱯᱨ	ᱯᱩ	ᱯᱪ	ᱯᱫ	ᱯᱬ	ᱯᱭ	ᱯᱮ	ᱯᱯ	/f/
23	KHI	ᱰᱨ	ᱰᱩ	ᱰᱪ	ᱰᱫ	ᱰᱬ	ᱰᱭ	ᱰᱮ	ᱰᱯ	/q/
24	KE	ᱱᱨ	ᱱᱩ	ᱱᱪ	ᱱᱫ	ᱱᱬ	ᱱᱭ	ᱱᱮ	ᱱᱯ	/k/
25	GE	ᱲᱨ	ᱲᱩ	ᱲᱪ	ᱲᱫ	ᱲᱬ	ᱲᱭ	ᱲᱮ	ᱲᱯ	/ᵑ/
26	NGI	ᱳᱨ	ᱳᱩ	ᱳᱪ	ᱳᱫ	ᱳᱬ	ᱳᱭ	ᱳᱮ	ᱳᱯ	/ŋ/
27	LE	ᱴᱨ	ᱴᱩ	ᱴᱪ	ᱴᱫ	ᱴᱬ	ᱴᱭ	ᱴᱮ	ᱴᱯ	/l/
28	ME	ᱵᱨ	ᱵᱩ	ᱵᱪ	ᱵᱫ	ᱵᱬ	ᱵᱭ	ᱵᱮ	ᱵᱯ	/m/
29	NE	ᱶᱨ	ᱶᱩ	ᱶᱪ	ᱶᱫ	ᱶᱬ	ᱶᱭ	ᱶᱮ	ᱶᱯ	/n/
30	HHE	ᱷᱨ	ᱷᱩ	ᱷᱪ	ᱷᱫ	ᱷᱬ	ᱷᱭ	ᱷᱮ	ᱷᱯ	/h/
31	VE	ᱸᱨ	ᱸᱩ	ᱸᱪ	ᱸᱫ	ᱸᱬ	ᱸᱭ	ᱸᱮ	ᱸᱯ	/v/
32	YE	ᱹᱨ	ᱹᱩ	ᱹᱪ	ᱹᱫ	ᱹᱬ	ᱹᱭ	ᱹᱮ	ᱹᱯ	/j/

10. Bibliography

- [1] The Unicode Standard Version 6.0.
- [2] ISO/IEC TR 15285 An operational model for characters and glyphs, First edition, 1998-12-15.
- [3] Yoshiki Mikami, A History of character Codes in Asia, 2002.
- [4] The Worlds Writing Systems, edited by P. Daniels, W. Bright, New York, Oxford University Press, 1996.
- [5] Kutadgu Bilig, Istanbul, Alaeddin Kiral Basimeve, 1942.
- [6] G. R. Rachmatl, Ergebnisse der Deutschen Turfan Forschung, 1936.

11. Appendix: Proposal Summary form

ISO/IEC JTC 1/SC 2/WG 2 PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646TP¹PT Please fill all the sections A, B and C below. Please read Principles and Procedures Document (P & P) from HTUhttp://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html ^{UT}H for guidelines and details before filling this form. Please ensure you are using the latest Form from HTUhttp://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html ^{UT}H. See also HTUhttp://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html ^{UT}H for latest Roadmaps.

A. Administrative

1. Title:	<i>Proposal for encoding the Turfan Mani script in the SMP of the UCS</i>
2. Requester's name:	<i>Nagaoka University of Technology</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Member body</i>
4. Submission date:	<i>2011-11-017</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>Yes</i>
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<i>Yes</i>
Proposed name of script:	<i>Uygur script</i>
b. The proposal is for addition of character(s) to an existing block:	<i>No</i>
Name of the existing block:	
2. Number of characters in proposal:	
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary	☐
B.1-Specialized (small collection)	☐
B.2-Specialized (large collection)	☐
C-Major extinct	☒
D-Attested extinct	☐
E-Minor extinct	☐
F-Archaic Hieroglyphic or Ideographic	☐
G-Obscure or questionable usage symbols	☐
4. Is a repertoire including character names provided?	<i>Yes</i>
a. If YES, are the names in accordance with the “character naming guidelines” In Annex L of P&P document?	<i>Yes</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>
5. Fonts Mani (TrueType)	

¹ Form number: N3902-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03)

related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Omarjan Osman</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Omerjan Osman, Information System and Control Engineering, Global Information Infrastructure Lab, Nagaoka University of Technology, Kamitomioka 1603-1, Nagaoka Shi, Niigata 940-2188, Japan, s075386@stn.nagaokaut.ac.jp</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>
7. Special encoding issues:	
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<i>Yes</i>
8. Additional Information:	
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.	

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	<i>No</i>
If YES explain	
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	<i>Yes</i>
If YES, with whom?	<i>Yoshiki Mikami Global Information Infrastructure Lab Nagaoka University of Technology Kamitomioka 1603-1, Nagaoka Shi, Niigata 940-2188, Japan. Tel: 090-1506-6595</i>

		<i>mikami@kjs.nagaokaut.ac.jp</i>
If YES, available relevant documents:	<i>Yes</i>	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?		
Reference:	<i>Design of Uyghur Character Information Processing System. 2009.3</i>	
4. The context of use for the proposed characters (type of use; common or rare)		
Reference:	<i>ISO/IEC JTC 1/SC 2/WG 2. 2008-10-23.</i>	
5. Are the proposed characters in current use by the user community?		
If YES, where?	Reference: <i>The objective of this paper is to create a standard character set for the Uyghur script, in order to help in studies on those ancient documents. Users in Germany, in USA, in Japan.</i>	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely		
In the BMP?	<i>No</i>	
If YES, is a rationale provided?		
If YES, reference:		
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		
		<i>Yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing Character or character sequence?		
		<i>No</i>
If YES, is a rationale for its inclusion provided?		
If YES, reference:		
9. Can any of the proposed characters be encoded using a composed character sequence of either Existing characters or other proposed characters?		
		<i>No</i>
If YES, is a rationale for its inclusion provided?		
If YES, reference:		
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?		
		<i>No</i>
If YES, is a rationale for its inclusion provided?		
If YES, reference:		
11. Does the proposal include use of combining characters and/or use of composite sequences?		
		<i>No</i>
If YES, is a rationale for such use provided?		
If YES, reference:		
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?		
If YES, reference:		
12. Does the proposal contain characters with any special properties such as		
		<i>No</i>

Control function or similar semantics?	
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	<i>No</i>
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	