



Internationalized Domain Names: Feedback of PAN L10n project on IDNAbis for Languages of Developing Asia¹

Version 1.01

*Sarmad Hussain and Nayyara Karamat
Center for Research in Urdu Language Processing
National University of Computer and Emerging Sciences
Lahore, Pakistan*

1. Introduction

This document presents the comparison of character set selection for internationalized domain names given in IDNAbis tables [1] and character by character choice done by local language teams under PAN localization project. Eight languages are considered namely Bengali, Dzongkha, Khmer, Lao, Mongolian, Nepali, Pashto and Urdu.

2. Label Separators

In IDNA2003 [2, 3, 4, 5], three characters are listed to be treated as label separators. In IDNAbis[1], only ASCII period is allowed to be used. If other characters are required for any language, they are expected to be mapped before using or storing the domain name. Language specific delimiters identified in our study are given in Table 2.1.

Table 2.1 List of Delimiters

Language	Character	Unicode
Dzongkha	ཤ	U+0F14
Urdu and Pashto	-	U+06D4

Language interface handling these particular languages will handle these delimiters in addition to FULL STOP U+002E. It is noted that all languages did not decide to use their sentence separator marks to be used as label separators. For example, In Nepali U+0964 is used as

¹ The language data discussed in this document has been provided by PAN Localization country partners. See Acknowledgements Section below.

sentence marker but they decided to use FULL STOP U+002E as label separator in domain names.

3. Comparison of Character Set Selection

Some characters are considered PVALID in IDNAbis table whereas not allowed in character by character decision of PAN teams. This case mostly occurred because a script is used to write multiple languages and there are many characters which are relevant for one language but not used in the other. Such characters can only be handled properly when language being used in the domain name is known. This language based filter can be applied on the top IDNAbis tables in language table at registry. More critical mismatches are where IDNAbis tables are disallowing some characters when in local community's opinion, they should be included. In this section, the comparison of character set selection using IDNAbis table and feedback from native speakers is presented.

3.1. Bengali

Bengali is an Indic language which uses Bengali script, closely related to Devanagiri script, both deriving from Brahmi script. Unicode block of Bengali is U+0980 – U+09FF. There are following PVALID characters which are either not used in Bengali or can be omitted without changing the meaning of word. So it is decided by Bengali community that these characters should be disallowed in Bengali domain names to reduce the risk of phishing. List of these characters is given in Table 3.1.

Table 3.1: Bengali Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
◌ে	0981	Bengali Sign Candrabindu	No	PVALID
◌ঌ	0982	Bengali Sign Anusvara	No	PVALID
◌্	09BC	Bengali Sign Nukta	No	PVALID
◌্ৰ	09BD	Bengali Sign Avagraha	No	PVALID
◌্ৰ্	09C4	Bengali Vowel Sign Vocalic RR	No	PVALID
◌্ৰ্ৰ	09E0	Bengali Letter Vocalic RR	No	PVALID
◌্ৰ্ৰ্	09E1	Bengali Letter Vocalic LL	No	PVALID
◌্ৰ্ৰ্ৰ	09E2	Bengali Vowel Sign Vocalic L	No	PVALID
◌্ৰ্ৰ্ৰ্	09E3	Bengali Vowel Sign Vocalic LL	No	PVALID
◌়া	09F0	Bengali Letter Ra With Middle Diagonal	No	PVALID
◌়া়	09F1	Bengali Letter Ra With Lower Diagonal	No	PVALID

In the Table 3.2, the Bengali characters are given which are disallowed by IDNAbis but required by the native speakers. These characters can be decomposed to PVALID characters with combination of Bengali Sign Nukta U+09BC but Bengali users are not accustomed with Nukta sign. It is rather disallowed in the table above to prevent ambiguous input. To solve this issue these characters will be decomposed at application level before referring to IDNAbis tables.

Table 3.2: Bengali Characters Requiring Decomposition

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]	Decomposed Form
ড়	09DC	Bengali letter RRA	Yes	DISALLOWED	09A1 09BC
ঢ়	09DD	Bengali letter RHA	Yes	DISALLOWED	09A2 09BC
য়	09DF	Bengali letter YYA	Yes	DISALLOWED	09AF 09BC

All the characters of Bengali block with their status decided by language community and IDNAbis algorithm are given in Appendix A.1.

3.2. Dzongkha

Dzongkha is written in Tibetan script which itself is motivated by the syllabic Devanagari writing system. Tibetan's Unicode block is U+0F00-U+0FFF. There are many characters which are allowed in IDNAbis table but in Dzongkha speakers view not likely to appear in the plain text. Such characters listed in Table 3.3 can be disallowed at registry level in language table.

Table 3.3: Dzongkha Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ཨྵ	0F00	TIBETAN SYLLABLE OM	NO	PVALID
ཨྶ	0F18	TIBETAN ASTROLOGICAL SIGN –KHYUD PA	NO	PVALID
ཨྷ	0F19	TIBETAN ASTROLOGICAL SIGN SDONG TSHUGS	NO	PVALID
ཨྸ	0F35	TIBETAN MARK NGAS BZUNG NYI ZLA	NO	PVALID
ཨྐྵ	0F37	TIBETAN MARK NGAS BZUNG SGOR RTAGS	NO	PVALID
ཨྺ	0F3E	TIBETAN SIGN YAR TSHES	NO	PVALID
ཨྻ	0F3F	TIBETAN SIGN MAR TSHES	NO	PVALID
ཨྼ	0F7E	TIBETAN SIGN RJES SU NGA RO	NO	PVALID
ཨ྽	0F82	TIBETAN SIGN NYI ZLA NAA DA	NO	PVALID
ཨ྿	0F83	TIBETAN SIGN SNA LDAN	NO	PVALID
ཨྺ	0F86	TIBETAN SIGN LCI RTAGS	NO	PVALID

འ	0F87	TIBETAN SIGN YANG RTAGS	NO	PVALID
མ	0F88	TIBETAN SIGN LCE TSA CAN	NO	PVALID
ཙ	0F89	TIBETAN SIGN MCHU CAN	NO	PVALID
ཚ	0F8A	TIBETAN SIGN GRU CAN RGYINGS	NO	PVALID
ཛ	0F8B	TIBETAN SIGN GRU MED RGYINGS	NO	PVALID
ཨ	0F96	TIBETAN SUBJOINED LETTER CHA	NO	PVALID
ཨ	0F9A	TIBETAN SUBJOINED LETTER TTA	NO	PVALID
ཨ	0F9B	TIBETAN SUBJOINED LETTER TTHA	NO	PVALID
ཨ	0F9C	TIBETAN SUBJOINED LETTER DDA	NO	PVALID
།	0FC6	TIBETAN SYMBOL PADMA GDAN	NO	PVALID

The characters listed in Table 3.4 are disallowed by IDNAbis table but required by language community. They have equivalent decomposed forms which are permitted in IDNAbis tables. So these characters can be used by user and proper normalization should be done at application level. At registry level, these characters will not be allowed.

Table 3.4: Dzongkha Characters Requiring Decomposition

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]	Decomposed Form
ཀྲ	0F43	TIBETAN LETTER GHHA	YES	DISALLOWED	0F42 0FB7
ཀླ	0F4D	TIBETAN LETTER DDHA	YES	DISALLOWED	0F4C 0FB7
ཀྴ	0F52	TIBETAN LETTER DHA	YES	DISALLOWED	0F51 0FB7
ཀྵ	0F57	TIBETAN LETTER BHA	YES	DISALLOWED	0F56 0FB7

འཇམ་	0F5C	TIBETAN LETTER DZHA	YES	DISALLOWED	0F5B 0FB7
ཀྲ	0F69	TIBETAN LETTER KSSA	YES	DISALLOWED	0F40 0FB5
འ	0F73	TIBETAN VOWEL SIGN II	YES	DISALLOWED	0F71 0F72
འ	0F75	TIBETAN VOWEL SIGN UU	YES	DISALLOWED	0F71 0F74
འཇམ་	0F93	TIBETAN SUBJOINED LETTER GHA	YES	DISALLOWED	0F92 0FB7
འཇམ་	0FA2	TIBETAN SUBJOINED LETTER DHA	YES	DISALLOWED	0FA1 0FB7
འཇམ་	0FA7	TIBETAN SUBJOINED LETTER BHA	YES	DISALLOWED	0FA6 0FB7
འཇམ་	0FAC	TIBETAN SUBJOINED LETTER DZHA	YES	DISALLOWED	0FAB 0FB7
འཇམ་	0FB9	TIBETAN SUBJOINED LETTER KSSA	YES	DISALLOWED	0F90 0FB5

There are two characters given in Table 3.5 which are used to mark syllable boundary in Dzongkha and are required to write words correctly. TIBETAN MARK INTERSYLLABIC TSHEG U+0F0B is required in domain names whereas TIBETAN MARK DELIMITER TSHEG BSTAR U+0F0C is recommended to be normalized to U+0F0B to avoid ambiguity in domain names. TIBETAN MARK INTERSYLLABIC TSHEG U+0F0B should be allowed in domain names.

Table 3.5: Dzongkha Characters Not Compatible with IDNAbis Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
.	0F0B	TIBETAN MARK INTERSYLLABIC TSHEG	YES	DISALLOWED
.	0F0C	TIBETAN MARK DELIMITER TSHEG BSTAR	YES but should be normalized to U+0F0B	DISALLOWED

Also refer to Table 2.1. above, which lists U+0F14 as Label Separator requirement for Dzongkha. This character must be DISALLOWED in the table for it to be usable in this context. Also, this character is confusable with U+0F7F, which is used in Sanskrit writing (used for Tibetan language and more rarely in Dzongkha). As it is confusable and rarely used in Dzongkha, Dzongkha community is currently considering blocking it in its language table at the ccTLD level.

All the characters of Dzongkha block with their status decided by language community and IDNAbis algorithm are given in Appendix A.2.

3.3. Khmer

Khmer is an Austro-Asiatic language and script and language of Cambodia. Its Unicode range is U+1780-U+17FF. In Table 3.6 the characters are listed which are not considered necessary to be used in Khmer domain names by native speakers. These characters can be restricted at registry level on top of the IDNAbis table.

Table 3.6: Khmer Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
៥	179D	KHMER LETTER SHA	NO	PVALID
៥	179E	KHMER LETTER SSO	NO	PVALID
អ	17A3	KHMER INDEPENDENT VOWEL QAQ	NO	PVALID
អ	17A4	KHMER INDEPENDENT VOWEL QAA	NO	PVALID
៊	17B4	KHMER VOWEL INHERENT AQ	NO	CONTEXTO
៊	17B5	KHMER VOWEL INHERENT AA	NO	CONTEXTO
័	17CE	KHMER SIGN KAKABAT	NO	PVALID
័	17D1	KHMER SIGN VIRIAM	NO	PVALID
័	17D3	KHMER SIGN BATHAMASAT	NO	PVALID
័	17DC	KHMER SIGN AVAKRAHASANYA	NO	PVALID
័	17DD	KHMER SIGN ATTHACAN	NO	PVALID

There is no character of Khmer found which is disallowed in IDNAbis but required by the language community of Khmer.

All the characters of Khmer block with their status decided by language community and IDNAbis algorithm are given in Appendix A.3.

3.4. Lao

Unicode code block for Lao is U+0E80–U+0EFF. Following are Lao characters in Table 3.7 which are not considered necessary for domain names by Lao community. These are allowed in IDNAbis table level but can be restricted at language table level.

Table 3.7: Lao Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ຸ	0EAF	LAO ELLIPSIS	NO	PVALID
ງ	0EC6	LAO KO LA	NO	PVALID

Following are the characters in Table 3.8 which require decomposition before using or storing the domain name.

Table 3.8: Lao Characters Requiring Decomposition

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]	Decomposed Form
ໝ	0EDC	LAO LETTER HO NO	Yes	DISALLOWED	0EAB 0E99
ໝ	0EDD	LAO LETTER HO MO	Yes	DISALLOWED	0EAB 0EA1

There is no character of Lao which is disallowed in IDNAbis but required by the language community of Lao. All the characters of Lao block with their status decided by language community and IDNAbis algorithm are given in Appendix A.4.

3.5. Mongolian

Mongolian is written in Cyrillic script. Cyrillic script has been derived from Greek script and has been traditionally used to write Slavic languages, including Russian. The Mongolian character set is slightly modified Cyrillic alphabet by adding two vowels such as (ø, γ). Unicode block of Cyrillic is U+0400 – U+04FF. There are many characters in Cyrillic script which are not used in Mongolian. Knowing the language, we can eliminate these characters from valid characters for Mongolian domain names. These characters are recommended to be excluded at registry level for Mongolian domain names. All such characters are listed in Table 3.8.

Table 3.8: Mongolian Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
è	450	CYRILLIC SMALL LETTER IE WITH GRAVE	NO	PVALID
Ђ	452	CYRILLIC SMALL LETTER DJE	NO	PVALID
Ѓ	453	CYRILLIC SMALL LETTER GJE	NO	PVALID
Є	454	CYRILLIC SMALL LETTER UKRAINIAN IE	NO	PVALID
ѕ	455	CYRILLIC SMALL LETTER DZE	NO	PVALID
і	456	CYRILLIC SMALL LETTER BYELORUSSIAN-UKRAINIANI	NO	PVALID
ї	457	CYRILLIC SMALL LETTER YI	NO	PVALID
ј	458	CYRILLIC SMALL LETTER JE	NO	PVALID
љ	459	CYRILLIC SMALL LETTER LJE	NO	PVALID
њ	45A	CYRILLIC SMALL LETTER NJE	NO	PVALID
ћ	45B	CYRILLIC SMALL	NO	PVALID

		LETTER TSHE		
ќ	45C	CYRILLIC SMALL LETTER KJE	NO	PVALID
ѝ	45D	CYRILLIC SMALL LETTER I WITH GRAVE	NO	PVALID
ѣ	45E	CYRILLIC SMALL LETTER SHORT U	NO	PVALID
џ	45F	CYRILLIC SMALL LETTER DZHE	NO	PVALID
Ѡ	461	CYRILLIC SMALL LETTER OMEGA	NO	PVALID
Ѣ	463	CYRILLIC SMALL LETTER YAT	NO	PVALID
Ѥ	465	CYRILLIC SMALL LETTER IOTIFIED E	NO	PVALID
Ѧ	467	CYRILLIC SMALL LETTER LITTLE YUS	NO	PVALID
Ѩ	469	CYRILLIC SMALL LETTER IOTIFIED LITTLE YUS	NO	PVALID
Ѭ	046B	CYRILLIC SMALL LETTER BIG YUS	NO	PVALID
Ѯ	046D	CYRILLIC SMALL LETTER IOTIFIED BIG YUS	NO	PVALID
Ѱ	046F	CYRILLIC SMALL LETTER KSI	NO	PVALID
Ѳ	471	CYRILLIC SMALL LETTER PSI	NO	PVALID
Ѵ	473	CYRILLIC SMALL LETTER FITA	NO	PVALID
Ѷ	475	CYRILLIC SMALL LETTER IZHITSA	NO	PVALID
Ѹ	477	CYRILLIC SMALL LETTER IZHITSA WITH DOUBLE GRAVE ACCENT	NO	PVALID
Ѻ	479	CYRILLIC SMALL LETTER UK	NO	PVALID
Ѽ	047B	CYRILLIC SMALL LETTER ROUND OMEGA	NO	PVALID
Ѿ	047D	CYRILLIC SMALL LETTER OMEGA WITH TITLO	NO	PVALID
ѿ	047F	CYRILLIC SMALL LETTER OT	NO	PVALID
Ҁ	481	CYRILLIC SMALL LETTER KOPPA	NO	PVALID

˘	483	COMBINING CYRILLIC TITLO	NO	CONTEXTO
˘	484	COMBINING CYRILLIC PALATALIZATION	NO	PVALID
˘	485	COMBINING CYRILLIC DASIA PNEUMATA	NO	PVALID
˘	486	COMBINING CYRILLIC PSILI PNEUMATA	NO	PVALID
й	048B	CYRILLIC SMALL LETTER SHORT I WITH TAIL	NO	PVALID
ь	048D	CYRILLIC SMALL LETTER SEMISOFT SIGN	NO	PVALID
ѣ	048F	CYRILLIC SMALL LETTER ER WITH TICK	NO	PVALID
ґ	491	CYRILLIC SMALL LETTER GHE WITH UPTURN (present in WGL4)	NO	PVALID
Ғ	493	CYRILLIC SMALL LETTER GHE WITH STROKE	NO	PVALID
Ҕ	495	CYRILLIC SMALL LETTER GHE WITH MIDDLE HOOK	NO	PVALID
Җ	497	CYRILLIC SMALL LETTER ZHE WITH DESCENDER	NO	PVALID
Ҙ	499	CYRILLIC SMALL LETTER ZE WITH DESCENDER	NO	PVALID
Қ	049B	CYRILLIC SMALL LETTER KA WITH DESCENDER	NO	PVALID
К	049D	CYRILLIC SMALL LETTER KA WITH VERTICAL STROKE	NO	PVALID
К	049F	CYRILLIC SMALL LETTER KA WITH STROKE	NO	PVALID
К	04A1	CYRILLIC SMALL LETTER BASHKIR KA	NO	PVALID
Ң	04A3	CYRILLIC SMALL LETTER EN WITH DESCENDER	NO	PVALID

Ғ	04A5	CYRILLIC SMALL LIGATURE EN GHE	NO	PVALID
Ѣ	04A7	CYRILLIC SMALL LETTER PE WITH MIDDLE HOOK	NO	PVALID
Ӏ	04A9	CYRILLIC SMALL LETTER ABKHASIAN HA	NO	PVALID
Ҕ	04AB	CYRILLIC SMALL LETTER ES WITH DESCENDER	NO	PVALID
Ҥ	04AD	CYRILLIC SMALL LETTER TE WITH DESCENDER	NO	PVALID
У	04AF	CYRILLIC SMALL LETTER STRAIGHT U	NO	PVALID
Ѹ	04B1	CYRILLIC SMALL LETTER STRAIGHT U WITH STROKE	NO	PVALID
Х	04B3	CYRILLIC SMALL LETTER HA WITH DESCENDER	NO	PVALID
Ц	04B5	CYRILLIC SMALL LIGATURE TE TSE	NO	PVALID
Ҷ	04B7	CYRILLIC SMALL LETTER CHE WITH DESCENDER	NO	PVALID
Ҧ	04B9	CYRILLIC SMALL LETTER CHE WITH VERTICAL STROKE	NO	PVALID
Һ	04BB	CYRILLIC SMALL LETTER SHHA	NO	PVALID
ӕ	04BD	CYRILLIC SMALL LETTER ABKHASIAN CHE	NO	PVALID
ӕ̣	04BF	CYRILLIC SMALL LETTER ABKHASIAN CHE WITH DESCENDER	NO	PVALID
Ӗ	04C2	CYRILLIC SMALL LETTER ZHE WITH BREVE	NO	PVALID
Ҝ	04C4	CYRILLIC SMALL LETTER KA WITH HOOK	NO	PVALID
Ҟ	04C6	CYRILLIC SMALL LETTER EL WITH	NO	PVALID

		TAIL		
Ң	04C8	CYRILLIC SMALL LETTER EN WITH HOOK	NO	PVALID
Ҥ	04CA	CYRILLIC SMALL LETTER EN WITH TAIL	NO	PVALID
Ҷ	04CC	CYRILLIC SMALL LETTER KHAKASSIAN CHE	NO	PVALID
Ҭ	04CE	CYRILLIC SMALL LETTER EM WITH TAIL	NO	PVALID
І	04CF	CYRILLIC SMALL LETTER PALOCHKA	NO	PVALID
ӑ	04D1	CYRILLIC SMALL LETTER A WITH BREVE	NO	PVALID
ӓ	04D3	CYRILLIC SMALL LETTER A WITH DIAERESIS	NO	PVALID
ӕ	04D5	CYRILLIC SMALL LIGATURE A IE	NO	PVALID
ӗ	04D7	CYRILLIC SMALL LETTER IE WITH BREVE	NO	PVALID
ә	04D9	CYRILLIC SMALL LETTER SCHWA	NO	PVALID
ӧ	04DB	CYRILLIC SMALL LETTER SCHWA WITH DIAERESIS	NO	PVALID
ӛ	04DD	CYRILLIC SMALL LETTER ZHE WITH DIAERESIS	NO	PVALID
ӟ	04DF	CYRILLIC SMALL LETTER ZE WITH DIAERESIS	NO	PVALID
Ӡ	4E1	CYRILLIC SMALL LETTER ABKHASIAN DZE	NO	PVALID
Ӣ	4E3	CYRILLIC SMALL LETTER I WITH MACRON	NO	PVALID
Ӥ	4E5	CYRILLIC SMALL LETTER I WITH DIAERESIS	NO	PVALID
Ӧ	4E7	CYRILLIC SMALL LETTER O WITH DIAERESIS	NO	PVALID
Ӱ	04EB	CYRILLIC SMALL LETTER BARRED O WITH	NO	PVALID

		DIAERESIS		
ë	04ED	CYRILLIC SMALL LETTER E WITH DIAERESIS	NO	PVALID
ÿ	04EF	CYRILLIC SMALL LETTER U WITH MACRON	NO	PVALID
ÿ	04F1	CYRILLIC SMALL LETTER U WITH DIAERESIS	NO	PVALID
ÿ̃	04F3	CYRILLIC SMALL LETTER U WITH DOUBLE ACUTE	NO	PVALID
č	04F5	CYRILLIC SMALL LETTER CHE WITH DIAERESIS	NO	PVALID
ґ	04F7	CYRILLIC SMALL LETTER GHE WITH DESCENDER	NO	PVALID
ѐ	04F9	CYRILLIC SMALL LETTER YERU WITH DIAERESIS	NO	PVALID
ґ̃	04FB	CYRILLIC SMALL LETTER GHE WITH STROKE AND HOOK	NO	PVALID
ҕ	04FD	CYRILLIC SMALL LETTER HA WITH HOOK	NO	PVALID
ҕ̃	04FF	CYRILLIC SMALL LETTER HA WITH STROKE	NO	PVALID

Cyrillic script requires case folding similar to Latin script. There are no characters found which are disallowed by IDNAbis and required by Mongolian users. All the characters of Cyrillic block with their status decided by language community and IDNAbis algorithm are given in Appendix A.5.

3.6. Nepali

Nepali is an Indo-Aryan language which uses Devanagari script derived from Brahmi script. Unicode block of Devanagari is U+0900 – U+097F. Nepali uses a subset of the script. Two consonants of Nepali which are not in popular used are recommended to be normalized to their more frequent variant forms. These characters should not be allowed at registry level. The variant are listed in Table 3.9.

Table 3.9: Nepali Character Variants

Character	Unicode	Normalized Character	Unicode of Normalized Character
न	929	𑑖	928
ळ	933	ल	932

Following characters of Devanagari script given in Table 3.10 are not used in Nepali. So these characters should be restricted at registry level.

Table 3.10: Nepali Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNabis Decision[1]
अे	0904	DEVANAGARI LETTER SHORT A	NO	PVALID
आँ	911	DEVANAGARI LETTER CANDRA O	NO	PVALID
औ	912	DEVANAGARI LETTER SHORT O	NO	PVALID
र	931	DEVANAGARI LETTER RRA	NO	PVALID
ळ	934	DEVANAGARI LETTER LLLA	NO	PVALID
े	944	DEVANAGARI VOWEL SIGN VOCALIC RR	NO	PVALID
े	945	DEVANAGARI VOWEL SIGN CANDRA E	NO	PVALID
े	946	DEVANAGARI VOWEL SIGN SHORT E	NO	PVALID
ो	949	DEVANAGARI VOWEL SIGN CANDRA O	NO	PVALID
ँ	951	DEVANAGARI STRESS SIGN UDATTA	NO	PVALID
ं	952	DEVANAGARI STRESS SIGN ANUDATTA	NO	PVALID
ँ	953	DEVANAGARI GRAVE ACCENT	NO	PVALID
ँ	954	DEVANAGARI ACUTE ACCENT	NO	PVALID
ऋ	960	DEVANAGARI LETTER VOCALIC RR	NO	PVALID
ॠ	961	DEVANAGARI LETTER VOCALIC LL	NO	PVALID
ॡ	962	DEVANAGARI VOWEL SIGN VOCALIC L	NO	PVALID
ॢ	963	DEVANAGARI VOWEL SIGN VOCALIC LL	NO	PVALID

ॠ	090C	DEVANAGARI LETTER VOCALIC L	NO	PVALID
ॡ	090D	DEVANAGARI LETTER CANDRA E	NO	PVALID
ॢ	090E	DEVANAGARI LETTER SHORT E	NO	PVALID
ॣ	093C	DEVANAGARI SIGN NUKTA	NO	PVALID
।	093D	DEVANAGARI SIGN AVAGRAHA	NO	PVALID
॥	094A	DEVANAGARI VOWEL SIGN SHORT O	NO	PVALID
१	97B	DEVANAGARI LETTER GGA	NO	PVALID
ॡ	97C	DEVANAGARI LETTER JJA	NO	PVALID
ॢ	97D	DEVANAGARI LETTER GLOTAL STOP	NO	PVALID
ॣ	97E	DEVANAGARI LETTER DDDA	NO	PVALID
।	97F	DEVANAGARI LETTER BBA	NO	PVALID

There is one character . U+970, DEVANAGARI ABBREVIATION SIGN which is required by Nepali users and not allowed in IDNAbis table. All the characters of Devanagari block with their status decided by language community and IDNAbis algorithm are given in Appendix A.6.

3.7. Pashto

Pashto is written in Arabic script and uses extended Arabic character set. The following Arabic block characters given in Table 3.11 are not used in plain text. These characters are not likely to be used in domain names for Pashto. It is recommended to exclude them from valid domain name characters to avoid confusing domain names.

Table 3.11: Pashto Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ـ	600	ARABIC NUMBER SIGN	NO	CONTEXT0
ـ	601	ARABIC SIGN SANAH	NO	CONTEXT0
ـ	601	ARABIC FOOTNOTE MARKER	NO	CONTEXT0
ـ	603	ARABIC SIGN SAFHA	NO	CONTEXT0
ـ	614	ARABIC SIGN TAKHALLUS	NO	PVALID
-	640	ARABIC TATWEEL	NO	PVALID

Pashto uses are subset of Arabic block and there are many characters which are not even known to Pashto speakers. Such characters should be restricted on Pashto registry level to avoid security risks. All such characters are listed Table 3.12.

Table 3.12: Arabic Script Characters not used in Pashto

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
◻	615	ARABIC SMALL HIGH TAH	NO	PVALID
ـ	657	ARABIC INVERTED DAMMA	NO	PVALID
◌◌◌	659	ARABIC ZWARAKAY	NO	PVALID
◌◌◌	065A	ARABIC VOWEL SIGN SMALL V ABOVE	NO	PVALID
◌◌◌	065B	ARABIC VOWEL SIGN INVERTED SMALL V ABOVE	NO	PVALID
◌◌◌	065C	ARABIC VOWEL SIGN DOT BELOW	NO	PVALID
◌◌◌	065D	ARABIC REVERSED DAMMA	NO	PVALID
◌◌◌	065E	ARABIC FATHA WITH TWO DOTS	NO	PVALID
ـ	066E	ARABIC LETTER DOTLESS BEH	NO	PVALID
ـ	066F	ARABIC LETTER DOTLESS QAF	NO	PVALID
ا	671	ARABIC LETTER ALEF WASLA	NO	PVALID
ا	672	ARABIC LETTER ALEF WITH WAVY HAMZA ABOVE	NO	PVALID

ا	673	ARABIC LETTER ALEF WITH WAVY HAMZA BELOW	NO	PVALID
ء	674	ARABIC LETTER HIGH HAMZA	NO	PVALID
آ	067A	ARABIC LETTER TTEHEH	NO	PVALID
أ	067B	ARABIC LETTER BEEH	NO	PVALID
آ	067D	ARABIC LETTER TEH WITH THREE DOTS ABOVE DOWN	NO	PVALID
آ	067F	ARABIC LETTER TEHEH	NO	PVALID
أ	680	ARABIC LETTER BEHEH	NO	PVALID
أ	681	ARABIC LETTER HAH WITH HAMZA ABOVE	NO	PVALID
أ	682	ARABIC LETTER HAH WITH TWO DOTS VERTICAL ABO	NO	PVALID
أ	683	ARABIC LETTER NYEH	NO	PVALID
أ	684	ARABIC LETTER DYEH	NO	PVALID
أ	687	ARABIC LETTER TCHEHEH	NO	PVALID
د	068A	ARABIC LETTER DAL WITH DOT BELOW	NO	PVALID
د	068B	ARABIC LETTER DAL WITH DOT BELOW AND SMALL T	NO	PVALID
د	068C	ARABIC LETTER DAHAL	NO	PVALID
د	068D	ARABIC LETTER DDAHAL	NO	PVALID
د	068E	ARABIC LETTER DUL	NO	PVALID
د	068F	ARABIC LETTER DAL WITH THREE DOTS ABOVE DOWN	NO	PVALID
د	690	ARABIC LETTER DAL WITH FOUR DOTS ABOVE	NO	PVALID
ر	692	ARABIC LETTER REH WITH SMALL V	NO	PVALID
ر	694	ARABIC LETTER REH WITH DOT BELOW	NO	PVALID
ر	695	ARABIC LETTER REH WITH SMALL V BELOW	NO	PVALID
ر	697	ARABIC LETTER REH WITH TWO DOTS ABOVE	NO	PVALID

ر	699	ARABIC LETTER REH WITH FOUR DOTS ABOVE	NO	PVALID
پ	069B	ARABIC LETTER SEEN WITH THREE DOTS BELOW	NO	PVALID
پ	069C	ARABIC LETTER SEEN WITH THREE DOTS BELOW AND	NO	PVALID
ص	069D	ARABIC LETTER SAD WITH TWO DOTS BELOW	NO	PVALID
ض	069E	ARABIC LETTER SAD WITH THREE DOTS ABOVE	NO	PVALID
ظ	069F	ARABIC LETTER TAH WITH THREE DOTS ABOVE	NO	PVALID
ع	06A0	ARABIC LETTER AIN WITH THREE DOTS ABOVE	NO	PVALID
ف	06A1	ARABIC LETTER DOTLESS FEH	NO	PVALID
ف	06A2	ARABIC LETTER FEH WITH DOT MOVED BELOW	NO	PVALID
ف	06A3	ARABIC LETTER FEH WITH DOT BELOW	NO	PVALID
ه	06A4	ARABIC LETTER VEH	NO	PVALID
ه	06A5	ARABIC LETTER FEH WITH THREE DOTS BELOW	NO	PVALID
ه	06A6	ARABIC LETTER PEHEH	NO	PVALID
ق	06A7	ARABIC LETTER QAF WITH DOT ABOVE	NO	PVALID
ق	06A8	ARABIC LETTER QAF WITH THREE DOTS ABOVE	NO	PVALID
ك	06AC	ARABIC LETTER KAF WITH DOT ABOVE	NO	PVALID
ك	06AD	ARABIC LETTER NG	NO	PVALID
ك	06AE	ARABIC LETTER KAF WITH THREE DOTS BELOW	NO	PVALID
گ	06B0	ARABIC LETTER GAF WITH RING	NO	PVALID
گ	06B1	ARABIC LETTER NGOEH	NO	PVALID
گ	06B2	ARABIC LETTER GAF WITH TWO DOTS BELOW	NO	PVALID
گ	06B3	ARABIC LETTER GUEH	NO	PVALID
گ	06B4	ARABIC LETTER GAF WITH THREE DOTS ABOVE	NO	PVALID

ل	06B5	ARABIC LETTER LAM WITH SMALL V	NO	PVALID
ل	06B6	ARABIC LETTER LAM WITH DOT ABOVE	NO	PVALID
ل	06B7	ARABIC LETTER LAM WITH THREE DOTS ABOVE	NO	PVALID
ل	06B8	ARABIC LETTER LAM WITH THREE DOTS BELOW	NO	PVALID
ن	06B9	ARABIC LETTER NOON WITH DOT BELOW	NO	PVALID
ن	06BA	ARABIC LETTER NOON GHUNNA	NO	PVALID
ن	06BB	ARABIC LETTER RNOON	NO	PVALID
ن	06BD	ARABIC LETTER NOON WITH THREE DOTS ABOVE	NO	PVALID
ن	06BF	ARABIC LETTER TCHEH WITH DOT ABOVE	NO	PVALID
و	06C4	ARABIC LETTER WAW WITH RING	NO	PVALID
و	06C5	ARABIC LETTER KIRGHIZ OE	NO	PVALID
و	06C6	ARABIC LETTER OE	NO	PVALID
و	06C7	ARABIC LETTER U	NO	PVALID
و	06C8	ARABIC LETTER YU	NO	PVALID
و	06C9	ARABIC LETTER KIRGHIZ YU	NO	PVALID
و	06CA	ARABIC LETTER WAW WITH TWO DOTS ABOVE	NO	PVALID
و	06CB	ARABIC LETTER VE	NO	PVALID
ي	06CE	ARABIC LETTER YEH WITH SMALL V	NO	PVALID
ي	06CF	ARABIC LETTER WAW WITH DOT ABOVE	NO	PVALID
ي	06D1	ARABIC LETTER YEH WITH THREE DOTS BELOW	NO	PVALID
ل	06D6	ARABIC SMALL HIGH LIGATURE SAD WITH LAM WITH	NO	PVALID
ل	06D7	ARABIC SMALL HIGH LIGATURE QAF WITH LAM WITH	NO	PVALID
ل	06D8	ARABIC SMALL HIGH MEEM	NO	PVALID

		INITIAL FORM		
ﻻ	06D9	ARABIC SMALL HIGH LAM ALEF	NO	PVALID
ﺞ	06DA	ARABIC SMALL HIGH JEEM	NO	PVALID
ﻻ	06DB	ARABIC SMALL HIGH THREE DOTS	NO	PVALID
ﺲ	06DC	ARABIC SMALL HIGH SEEN	NO	PVALID
ﻻ	06DD	ARABIC END OF AYAH	NO	CONTEXTO
ﻻ	06DF	ARABIC SMALL HIGH ROUNDED ZERO	NO	PVALID
ﻻ	06E0	ARABIC SMALL HIGH UPRIGHT RECTANGULAR ZERO	NO	PVALID
ﻻ	06E1	ARABIC SMALL HIGH DOTLESS HEAD OF KHAH	NO	PVALID
ﻻ	06E2	ARABIC SMALL HIGH MEEM ISOLATED FORM	NO	PVALID
ﺲ	06E3	ARABIC SMALL LOW SEEN	NO	PVALID
ﻻ	06E4	ARABIC SMALL HIGH MADDA	NO	PVALID
ﻻ	06E5	ARABIC SMALL WAW	NO	PVALID
ﻻ	06E6	ARABIC SMALL YEY	NO	PVALID
ﻻ	06E7	ARABIC SMALL HIGH YEY	NO	PVALID
ﻻ	06E8	ARABIC SMALL HIGH NOON	NO	PVALID
ﻻ	06EA	ARABIC EMPTY CENTRE LOW STOP	NO	PVALID
ﻻ	06EB	ARABIC EMPTY CENTRE HIGH STOP	NO	PVALID
ﻻ	06EC	ARABIC ROUNDED HIGH STOP WITH FILLED CENTRE	NO	PVALID
ﻻ	06ED	ARABIC SMALL LOW MEEM	NO	PVALID
ﺩ	06EE	ARABIC LETTER DAL WITH INVERTED V	NO	PVALID
ﺭ	06EF	ARABIC LETTER REH WITH INVERTED V	NO	PVALID
ﺶ	06FA	ARABIC LETTER SHEEN WITH DOT BELOW	NO	PVALID
ﺾ	06FB	ARABIC LETTER DAD WITH DOT BELOW	NO	PVALID
ﻐ	06FC	ARABIC LETTER GHAIN WITH DOT BELOW	NO	PVALID
ﻩ	06FF	ARABIC LETTER HEH WITH INVERTED V	NO	PVALID

In Arabic character set there are many such characters which are repeated for different language allowing redundancy. For example, two sets of digits are there, one for Arabic language and other

for other language e.g. Persian, Urdu. There is very likely chance that users will confuse between these characters. It is suggested to normalize these characters on look up. Following are such characters and their normalized characters for Pashto.

Table 3.13: Normalization in Pashto

Character	Unicode	Normalized Character	Unicode of Normalized Character
ک	643	ک	06A9
ښ	679	ښ	067C
ډ	688	ډ	689
ژ	691	ژ	0693
ک	06AA	ک	06A9
گ	06AB	گ	06AF
ه	06BE	ه	0647
ه	06C1	ه	0647
ه	06C2	ه	06C0
ه	06C3	ه	0629
۱	06F1	۱	0661
۲	06F2	۲	0662
۳	06F3	۳	0663
۴	06F4	۴	0664
۵	06F5	۵	0665
۶	06F6	۶	0666
۷	06F7	۷	0667
۸	06F8	۸	0668
۹	06F9	۹	0669

ى	06CC	ى	0649
ے	06D2	ى	0649
ئ	06D3	ى	0626
ي	064A	ى	0649
ا	0625	ا	0627

All the characters of Arabic block with their status decided by language community and IDNAbis algorithm are given in Appendix A.7.

3.8. Sinhala

Sinhala is an Indo-Aryan language, spoken in Sri Lanka by 13 million people and also know as Sinhalese and Singhalese [7]. Sinhala is one of the national languages of Sri Lanka. It has significantly different written and spoken form, with liberty from influenced by Sanskrit [8, 6]. Sinhala script is descendant of Brahmi script and was formed between 7th -8th century [6] and has a syllabic written system, like other Indic scripts. However the system is unique because it has distinct rounded forms with no top-line, similar to south Indian scripts, latter used to write Dravidian languages.

In Sinhala only one character was not supported by the language community and it was also DISALLOWED by the IDNAbis process. All remaining characters that are required by the language community are also declared PVALID by the IDNAbis process.

All the characters Sinhala block with their status decided by language community and IDNAbis algorithm are given in Appendix A.8.

3.9. Urdu

Urdu is written in Arabic script and uses extended Arabic character set similarly as Pashto. As shown in Table 3.11 for Pashto, following characters given in Table 3.14 are also not recommended to be used in Urdu domain names.

Table 3.14: Urdu Characters Disallowed in Language Table

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
۰	600	ARABIC NUMBER SIGN	NO	CONTEXT0
۱	601	ARABIC SIGN SANAH	NO	CONTEXT0
۲	601	ARABIC FOOTNOTE MARKER	NO	CONTEXT0
۳	603	ARABIC SIGN SAFHA	NO	CONTEXT0

ٸ	614	ARABIC SIGN TAKHALLUS	NO	PVALID
-	640	ARABIC TATWEEL	NO	PVALID

Following are the Arabic block characters which are not used in Urdu. At registry level these characters should not be allowed in Urdu domain names.

Table 3.15: Arabic Script Characters not used in Urdu

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
◻	615	ARABIC SMALL HIGH TAH	NO	PVALID
.	657	ARABIC INVERTED DAMMA	NO	PVALID
◌◌◌	659	ARABIC ZWARAKAY	NO	PVALID
◌◌◌	065A	ARABIC VOWEL SIGN SMALL V ABOVE	NO	PVALID
◌◌◌	065B	ARABIC VOWEL SIGN INVERTED SMALL V ABOVE	NO	PVALID
◌◌◌	065C	ARABIC VOWEL SIGN DOT BELOW	NO	PVALID
◌◌◌	065D	ARABIC REVERSED DAMMA	NO	PVALID
◌◌◌	065E	ARABIC FATHA WITH TWO DOTS	NO	PVALID
◌◌◌	066E	ARABIC LETTER DOTLESS BEH	NO	PVALID
◌◌◌	066F	ARABIC LETTER DOTLESS QAF	NO	PVALID
اَ	671	ARABIC LETTER ALEF WASLA	NO	PVALID
اِ	672	ARABIC LETTER ALEF WITH WAVY HAMZA ABOVE	NO	PVALID
اِ	673	ARABIC LETTER ALEF WITH WAVY HAMZA BELOW	NO	PVALID
ء	674	ARABIC LETTER HIGH HAMZA	NO	PVALID
ح	067A	ARABIC LETTER TTEHEH	NO	PVALID
ح	067B	ARABIC LETTER BEEH	NO	PVALID
ح	067C	ARABIC LETTER TEH WITH RING	NO	PVALID
ح	067D	ARABIC LETTER TEH WITH THREE DOTS ABOVE DOWN	NO	PVALID
ح	067F	ARABIC LETTER TEHEH	NO	PVALID
ح	680	ARABIC LETTER BEHEH	NO	PVALID
ح	681	ARABIC LETTER HAH WITH HAMZA ABOVE	NO	PVALID
ح	682	ARABIC LETTER HAH WITH TWO DOTS VERTICAL ABO	NO	PVALID
ح	683	ARABIC LETTER NYEH	NO	PVALID
ح	684	ARABIC LETTER DYEH	NO	PVALID
ح	685	ARABIC LETTER HAH WITH THREE DOTS ABOVE	NO	PVALID

جھ	687	ARABIC LETTER TCHEHEH	NO	PVALID
دھ	689	ARABIC LETTER DAL WITH RING	NO	PVALID
دٓ	068A	ARABIC LETTER DAL WITH DOT BELOW	NO	PVALID
دٓٓ	068B	ARABIC LETTER DAL WITH DOT BELOW AND SMALL T	NO	PVALID
ذ	068C	ARABIC LETTER DAHAL	NO	PVALID
ذٓ	068D	ARABIC LETTER DDAHAL	NO	PVALID
ذٓٓ	068E	ARABIC LETTER DUL	NO	PVALID
ذٓٓٓ	068F	ARABIC LETTER DAL WITH THREE DOTS ABOVE DOWN	NO	PVALID
ڙ	690	ARABIC LETTER DAL WITH FOUR DOTS ABOVE	NO	PVALID
ڙٓ	692	ARABIC LETTER REH WITH SMALL V	NO	PVALID
ڙٓٓ	693	ARABIC LETTER REH WITH RING	NO	PVALID
ڙٓٓٓ	694	ARABIC LETTER REH WITH DOT BELOW	NO	PVALID
ڙٓٓٓٓ	695	ARABIC LETTER REH WITH SMALL V BELOW	NO	PVALID
ڙٓٓٓٓٓ	696	ARABIC LETTER REH WITH DOT BELOW AND DOT ABO	NO	PVALID
ڙٓٓٓٓٓٓ	697	ARABIC LETTER REH WITH TWO DOTS ABOVE	NO	PVALID
ڙٓٓٓٓٓٓٓ	699	ARABIC LETTER REH WITH FOUR DOTS ABOVE	NO	PVALID
سٓ	069A	ARABIC LETTER SEEN WITH DOT BELOW AND DOT AB	NO	PVALID
سٓٓ	069B	ARABIC LETTER SEEN WITH THREE DOTS BELOW	NO	PVALID
سٓٓٓ	069C	ARABIC LETTER SEEN WITH THREE DOTS BELOW AND	NO	PVALID
سٓٓٓٓ	069D	ARABIC LETTER SAD WITH TWO DOTS BELOW	NO	PVALID
سٓٓٓٓٓ	069E	ARABIC LETTER SAD WITH THREE DOTS ABOVE	NO	PVALID
ظ	069F	ARABIC LETTER TAH WITH THREE DOTS ABOVE	NO	PVALID
عٓٓٓ	06A0	ARABIC LETTER AIN WITH THREE DOTS ABOVE	NO	PVALID
وٓ	06A1	ARABIC LETTER DOTLESS FEH	NO	PVALID
وٓٓ	06A2	ARABIC LETTER FEH WITH DOT MOVED BELOW	NO	PVALID
وٓٓٓ	06A3	ARABIC LETTER FEH WITH DOT BELOW	NO	PVALID
وٓٓٓٓ	06A4	ARABIC LETTER VEH	NO	PVALID
وٓٓٓٓٓ	06A5	ARABIC LETTER FEH WITH THREE DOTS BELOW	NO	PVALID
قٓ	06A6	ARABIC LETTER PEHEH	NO	PVALID
قٓٓ	06A7	ARABIC LETTER QAF WITH DOT ABOVE	NO	PVALID
قٓٓٓ	06A8	ARABIC LETTER QAF WITH THREE DOTS ABOVE	NO	PVALID
ڪا	06AA	ARABIC LETTER SWASH KAF	NO	PVALID

ك	06AB	ARABIC LETTER KAF WITH RING	NO	PVALID
ك	06AC	ARABIC LETTER KAF WITH DOT ABOVE	NO	PVALID
ك	06AD	ARABIC LETTER NG	NO	PVALID
ك	06AE	ARABIC LETTER KAF WITH THREE DOTS BELOW	NO	PVALID
ك	06B0	ARABIC LETTER GAF WITH RING	NO	PVALID
ك	06B1	ARABIC LETTER NGOEH	NO	PVALID
ك	06B2	ARABIC LETTER GAF WITH TWO DOTS BELOW	NO	PVALID
ك	06B3	ARABIC LETTER GUEH	NO	PVALID
ك	06B4	ARABIC LETTER GAF WITH THREE DOTS ABOVE	NO	PVALID
ل	06B5	ARABIC LETTER LAM WITH SMALL V	NO	PVALID
ل	06B6	ARABIC LETTER LAM WITH DOT ABOVE	NO	PVALID
ل	06B7	ARABIC LETTER LAM WITH THREE DOTS ABOVE	NO	PVALID
ل	06B8	ARABIC LETTER LAM WITH THREE DOTS BELOW	NO	PVALID
ن	06B9	ARABIC LETTER NOON WITH DOT BELOW	NO	PVALID
ن	06BB	ARABIC LETTER RNOON	NO	PVALID
ن	06BC	ARABIC LETTER NOON WITH RING	NO	PVALID
ن	06BD	ARABIC LETTER NOON WITH THREE DOTS ABOVE	NO	PVALID
ن	06BF	ARABIC LETTER TCHEH WITH DOT ABOVE	NO	PVALID
و	06C4	ARABIC LETTER WAW WITH RING	NO	PVALID
و	06C5	ARABIC LETTER KIRGHIZ OE	NO	PVALID
و	06C6	ARABIC LETTER OE	NO	PVALID
و	06C7	ARABIC LETTER U	NO	PVALID
و	06C8	ARABIC LETTER YU	NO	PVALID
و	06C9	ARABIC LETTER KIRGHIZ YU	NO	PVALID
و	06CA	ARABIC LETTER WAW WITH TWO DOTS ABOVE	NO	PVALID
و	06CB	ARABIC LETTER VE	NO	PVALID
ي	06CD	ARABIC LETTER YEH WITH TAIL	NO	PVALID
ي	06CE	ARABIC LETTER YEH WITH SMALL V	NO	PVALID
ي	06CF	ARABIC LETTER WAW WITH DOT ABOVE	NO	PVALID
ي	06D0	ARABIC LETTER E	NO	PVALID
ي	06D1	ARABIC LETTER YEH WITH THREE DOTS BELOW	NO	PVALID
ه	06D5	ARABIC LETTER AE	NO	PVALID
ط	06D6	ARABIC SMALL HIGH LIGATURE SAD WITH LAM WITH	NO	PVALID
ظ	06D7	ARABIC SMALL HIGH LIGATURE QAF WITH LAM WITH	NO	PVALID
م	06D8	ARABIC SMALL HIGH MEEM INITIAL FORM	NO	PVALID
ل	06D9	ARABIC SMALL HIGH LAM ALEF	NO	PVALID

ج	06DA	ARABIC SMALL HIGH JEEM	NO	PVALID
ٴ	06DB	ARABIC SMALL HIGH THREE DOTS	NO	PVALID
س	06DC	ARABIC SMALL HIGH SEEN	NO	PVALID
ٴ	06DF	ARABIC SMALL HIGH ROUNDED ZERO	NO	PVALID
ٴ	06E0	ARABIC SMALL HIGH UPRIGHT RECTANGULAR ZERO	NO	PVALID
ٴ	06E1	ARABIC SMALL HIGH DOTLESS HEAD OF KHAH	NO	PVALID
م	06E2	ARABIC SMALL HIGH MEEM ISOLATED FORM	NO	PVALID
س	06E3	ARABIC SMALL LOW SEEN	NO	PVALID
ٴ	06E4	ARABIC SMALL HIGH MADDA	NO	PVALID
و	06E5	ARABIC SMALL WAW	NO	PVALID
ٴ	06E6	ARABIC SMALL YEH	NO	PVALID
ٴ	06E7	ARABIC SMALL HIGH YEH	NO	PVALID
ن	06E8	ARABIC SMALL HIGH NOON	NO	PVALID
ٴ	06EA	ARABIC EMPTY CENTRE LOW STOP	NO	PVALID
ٴ	06EB	ARABIC EMPTY CENTRE HIGH STOP	NO	PVALID
ٴ	06EC	ARABIC ROUNDED HIGH STOP WITH FILLED CENTRE	NO	PVALID
م	06ED	ARABIC SMALL LOW MEEM	NO	PVALID
ذ	06EE	ARABIC LETTER DAL WITH INVERTED V	NO	PVALID
ر	06EF	ARABIC LETTER REH WITH INVERTED V	NO	PVALID
ثب	06FA	ARABIC LETTER SHEEN WITH DOT BELOW	NO	PVALID
ظب	06FB	ARABIC LETTER DAD WITH DOT BELOW	NO	PVALID
غ	06FC	ARABIC LETTER GHAIN WITH DOT BELOW	NO	PVALID
ه	06FF	ARABIC LETTER HEH WITH INVERTED V	NO	PVALID

There are characters with the similar shape which are confusable with each other and each language makes its own decision for which one to include and others to normalize to that character. Such characters and their normalized characters for Urdu are listed in Table 3.16.

Table 3.17: Normalization in Urdu

Character	Unicode	Normalized Character	Unicode of Normalized character
ا	625	ا	627
ئ	626	ء	621
ة	629	ة	6C3

ك	643	ک	06A9
ه	647	ہ	6C1
ی	649	ی	06CC
۰	660	۰	06F0
۱	661	۱	06F1
۲	662	۲	06F2
۳	663	۳	06F3
۴	664	۴	06F4
۵	665	۵	06F5
۶	666	۶	06F6
۷	667	۷	06F7
۸	668	۸	06F8
۹	669	۹	06F9
ي	064A	ی	06CC
۰	06C0	۰	06C2

All the characters of Arabic block with their status decided by language community and IDNabis algorithm are given in Appendix A.9.

4. Discussion

Urdu and Pashto are written in the same script i.e. Arabic script which is cursive in nature. It does not have concept of space between words, rather final form of character at the end indicates word boundary. In English if we want to write multiple word url we can separate it using hyphen, 'two-word' or even if we simply write them together they are easily readable, 'twowords'. In Urdu and Pashto if we write words without any separator character, they may assume wrong shape decreasing the readability. For example, 'دس دن' (ten days) will be display erroneously 'دسدن' if written without separator. The final shape of character can be obtained by putting space or zero width non-joiner at the end of word. Users are more familiar with use of space instead of zero width non-joiner. Majority of the users even don't know about the concept of ZWNJ. To cater this issue, it is recommended that space should be allowed in domain names at interface level and should be replaced with ZWNJ before processing.

Arabic script include multiple code points which have similar display shape and it is language specific decision which code point is relevant to it. If multiple code points having same shape are

allowed in a language it can be a source of phishing. For example, Arabic letters ◌ U+647 and ◌ U+06C1 have the same shape. First one is used in Arabic, Pashto and some other languages whereas second is used in Urdu. To prevent the security risk, such characters are normalized according to the Table 3.11 and Table 3.17 given in Pashto and Urdu sections.

In traditional domain name system, country level domain names are assigned. In the new system of internationalized domain names, it is not sufficient. There can be multiple languages in a country or a single language being spoken across the boundaries of countries. A single script is also used for many languages. For example, there are multiple languages in Pakistan which use Arabic script. Character sets and normalization rules for each language are different. To handle all the languages, sub domains can be maintained under country level domain name.

5. Acknowledgements

The data for the tables has been provided by the country teams of PAN Localization project in Afghanistan, Bangladesh, Bhutan, Cambodia, Laos, Mongolia, Nepal and Pakistan. Details of these teams and memberships are available at <http://www.pan10n.net/english/partners.htm>. The project has been funded by PAN Localization Project grant by International Development Research Center, Canada (www.IDRC.ca).

6. References

- [1] <http://stupid.domain.name/idnabis/>
- [2] P. Hoffman, and M. Blanchet, "Preparation of Internationalized Strings ("stringprep")" <http://www.rfc-editor.org/rfc/rfc3454.txt>, 2002.
- [3] P. Faltstrom, P. Hoffman, and A. Costello, "Internationalizing Domain Names in Applications (IDNA)" <http://www.rfc-editor.org/rfc/rfc3490.txt>, 2003.
- [4] P. Hoffman, M. B. Viagenie, "Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)" <http://www.rfc-editor.org/rfc/rfc3491.txt>, 2003.
- [5] A. Costello, "Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)," <http://www.rfc-editor.org/rfc/rfc3492.txt>, 2003.
- [6] Coulmas, F. (1996). *Encyclopedia of Writing Systems*. Blackwell Publishers, Oxford, UK.
- [7] Ethnologue.com. "Sinhala: A language of Sri Lanka." Retrieved from http://www.ethnologue.com/show_language.asp?code=sin on 26th April, 2007.
- [8] Disanayaka, J. B. (2003). *Say it in Sinhala*. Stamford Lake, Pannipitiya, Sri Lanka.

Appendices

Appendix A.1: Bengali Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ঁ	0981	Bengali Sign Candrabindu	No	PVALID
ং	0982	Bengali Sign Anusvara	No	PVALID
ঃ	0983	Bengali Sign Visarga	YES	PVALID
	0984	<reserved>		UNASSIGNED
অ	0985	Bengali letter A	YES	PVALID
আ	0986	Bengali letter AA	YES	PVALID
ই	0987	Bengali letter I	YES	PVALID
ঈ	0988	Bengali letter II	YES	PVALID
উ	0989	Bengali letter U	YES	PVALID
ঊ	098A	Bengali letter UU	YES	PVALID
ঋ	098B	Bengali letter Vocalic R	YES	PVALID
ৠ	098C	Bengali letter Vocalic L	YES	PVALID
	098D	<reserved>		UNASSIGNED
	098E	<reserved>		UNASSIGNED
এ	098F	Bengali letter E	YES	PVALID
ঐ	0990	Bengali letter AI	YES	PVALID
	0991	<reserved>		UNASSIGNED
	0992	<reserved>		UNASSIGNED
ও	0993	Bengali letter O	YES	PVALID
ঔ	0994	Bengali letter AU	YES	PVALID
ক	0995	Bengali letter KA	YES	PVALID
খ	0996	Bengali letter KHA	YES	PVALID
গ	0997	Bengali letter GA	YES	PVALID
ঘ	0998	Bengali letter GHA	YES	PVALID
ঙ	0999	Bengali letter NGA	YES	PVALID
চ	099A	Bengali letter CA	YES	PVALID
ছ	099B	Bengali letter CHA	YES	PVALID
জ	099C	Bengali letter JA	YES	PVALID
ঝ	099D	Bengali letter JHA	YES	PVALID
ঞ	099E	Bengali letter NYA	YES	PVALID
ট	099F	Bengali letter TTA	YES	PVALID
ঠ	09A0	Bengali letter TTAH	YES	PVALID
ড	09A1	Bengali letter DDA	YES	PVALID
ঢ	09A2	Bengali letter DDHA	YES	PVALID
ণ	09A3	Bengali letter NNA	YES	PVALID
ত	09A4	Bengali letter TA	YES	PVALID
থ	09A5	Bengali letter THA	YES	PVALID
দ	09A6	Bengali letter DA	YES	PVALID

ধ	09A7	Bengali letter DAH	YES	PVALID
ন	09A8	Bengali letter NA	YES	PVALID
	09A9	<reserved>		UNASSIGNED
প	09AA	Bengali letter PA	YES	PVALID
ফ	09AB	Bengali letter PHA	YES	PVALID
ব	09AC	Bengali letter BA	YES	PVALID
ভ	09AD	Bengali letter BHA	YES	PVALID
ম	09AE	Bengali letter MA	YES	PVALID
য	09AF	Bengali letter YA	YES	PVALID
র	09B0	Bengali letter RA	YES	PVALID
	09B1	<reserved>		UNASSIGNED
ল	09B2	Bengali letter LA	YES	PVALID
	09B3	<reserved>		UNASSIGNED
	09B4	<reserved>		UNASSIGNED
	09B5	<reserved>		UNASSIGNED
শ	09B6	Bengali letter SHA	YES	PVALID
ষ	09B7	Bengali letter SSA	YES	PVALID
স	09B8	Bengali letter SA	YES	PVALID
হ	09B9	Bengali letter HA	YES	PVALID
	09BA	<reserved>		UNASSIGNED
	09BB	<reserved>		UNASSIGNED
্	09BC	Bengali Sign Nukta	No	PVALID
্	09BD	Bengali Sign Avagraha	No	PVALID
া	09BE	Bengali Vowel Sign AA	YES	PVALID
ি	09BF	Bengali Vowel Sign I	YES	PVALID
ী	09C0	Bengali Vowel Sign II	YES	PVALID
ু	09C1	Bengali Vowel Sign U	YES	PVALID
ূ	09C2	Bengali Vowel Sign UU	YES	PVALID
্	09C3	Bengali Vowel Sign R	YES	PVALID
্	09C4	Bengali Vowel Sign Vocalic RR	No	PVALID
	09C5	<reserved>		UNASSIGNED
	09C6	<reserved>		UNASSIGNED
ে	09C7	Bengali Vowel Sign E	YES	PVALID
ৈ	09C8	Bengali Vowel Sign AI	YES	PVALID
	09C9	<reserved>		UNASSIGNED
	09CA	<reserved>		UNASSIGNED
ো	09CB	Bengali Vowel Sign O	YES	PVALID
ৌ	09CC	Bengali Vowel Sign AU	YES	PVALID
্	09CD	Bengali Sign Virama (hasant)	YES	PVALID
ৎ	09CE	Bengali letter Khanda Ta	No	PVALID
	09CF	<reserved>		UNASSIGNED
	09D0	<reserved>		UNASSIGNED
	09D1	<reserved>		UNASSIGNED
	09D2	<reserved>		UNASSIGNED

	09D3	<reserved>		UNASSIGNED
	09D4	<reserved>		UNASSIGNED
	09D5	<reserved>		UNASSIGNED
	09D6	<reserved>		UNASSIGNED
ী	09D7	Bengali AU Length Mark	YES	PVALID
	09D8	<reserved>		UNASSIGNED
	09D9	<reserved>		UNASSIGNED
	09DA	<reserved>		UNASSIGNED
	09DB	<reserved>		UNASSIGNED
ড়	09DC	Bengali letter RRA	YES	DISALLOWED
ঢ	09DD	Bengali letter RHA	YES	DISALLOWED
	09DE	<reserved>		UNASSIGNED
য়	09DF	Bengali letter YYA	YES	DISALLOWED
ঋ	09E0	Bengali Letter Vocalic RR	No	PVALID
ৠ	09E1	Bengali Letter Vocalic LL	No	PVALID
ঌ	09E2	Bengali Vowel Sign Vocalic L	No	PVALID
঍	09E3	Bengali Vowel Sign Vocalic LL	No	PVALID
	09E4	<reserved>		UNASSIGNED
	09E5	<reserved>		UNASSIGNED
০	09E6	Bengali Digit Zero	YES	PVALID
১	09E7	Bengali Digit One	YES	PVALID
২	09E8	Bengali Digit Two	YES	PVALID
৩	09E9	Bengali Digit Three	YES	PVALID
৪	09EA	Bengali Digit Four	YES	PVALID
৫	09EB	Bengali Digit Five	YES	PVALID
৬	09EC	Bengali Digit Six	YES	PVALID
৭	09ED	Bengali Digit Seven	YES	PVALID
৮	09EE	Bengali Digit Eight	YES	PVALID
৯	09EF	Bengali Digit Nine	YES	PVALID
৺	09F0	Bengali Letter Ra With Middle Diagonal	No	PVALID
৻	09F1	Bengali Letter Ra With Lower Diagonal	No	PVALID
ৼ	09F2	Bengali Rupee Mark	No	DISALLOWED
৽	09F3	Bengali Rupee Sign	No	DISALLOWED
৾	09F4	Bengali Currency Numerator One	No	DISALLOWED
৿	09F5	Bengali Currency Numerator Two	No	DISALLOWED
৺	09F6	Bengali Currency Numerator Three	No	DISALLOWED
৻	09F7	Bengali Currency Numerator Four	No	DISALLOWED
ৼ	09F8	Bengali Currency Numerator One Less Than The Denominator	No	DISALLOWED
৽	09F9	Bengali Currency Denominator Sixteen	No	DISALLOWED

৯	09FA	Bengali Isshar	No	DISALLOWED
	200C	Zero Width Non Joiner	No	CONTEXTJ
	200D	Zero Width Joiner	No	CONTEXTJ

Appendix A.2: Dzongkha Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ཨ	0F00	TIBETAN SYLLABLE OM	NO	PVALID
ཨ	0F01	TIBETAN MARK GTER YIG MGO TRUNCATED A	NO	DISALLOWED
ཨུམ་བུ་	0F02	TIBETAN MARK GTER YIG MGO -UM RNAM BCAD MA	NO	DISALLOWED
ཨུམ་བུ་	0F03	TIBETAN MARK GTER YIG MGO -UM GTER TSHEG MA	NO	DISALLOWED
ཨུམ་བུ་	0F04	TIBETAN MARK INITIAL YIG MGO MDUN MA	NO	DISALLOWED
ཨུམ་བུ་	0F05	TIBETAN MARK CLOSING YIG MGO SGAB MA	NO	DISALLOWED
ཨུམ་བུ་	0F06	TIBETAN MARK CARET YIG MGO PHUR SHAD MA	NO	DISALLOWED
ཨུམ་བུ་	0F07	TIBETAN MARK YIG MGO TSHEG SHAD MA	NO	DISALLOWED
ཨུམ་བུ་	0F08	TIBETAN MARK SBRUL SHAD	NO	DISALLOWED
ཨུམ་བུ་	0F09	TIBETAN MARK BSKUR YIG MGO	NO	DISALLOWED
ཨུམ་བུ་	0F0A	TIBETAN MARK BKA- SHOG YIG MGO	NO	DISALLOWED
.	0F0B	TIBETAN MARK INTERSYLLABIC TSHEG	YES	DISALLOWED
.	0F0C	TIBETAN MARK DELIMITER TSHEG BSTAR	YES (variant of base character)	DISALLOWED
།	0F0D	TIBETAN MARK SHAD	NO	DISALLOWED
།།	0F0E	TIBETAN MARK NYIS SHAD	NO	DISALLOWED
།	0F0F	TIBETAN MARK TSHEG SHAD	NO	DISALLOWED
།	0F10	TIBETAN MARK NYIS TSHEG SHAD	NO	DISALLOWED

ཨ	0F11	TIBETAN MARK RIN CHEN SPUNGS SHAD	NO	DISALLOWED
ཨྱ	0F12	TIBETAN MARK RGYA GRAM SHAD	NO	DISALLOWED
ཨྲ	0F13	TIBETAN MARK CARET -DZUD RTAGS ME LONG CAN	NO	DISALLOWED
ཨླ	0F14	TIBETAN MARK GTER TSHEG	NO	DISALLOWED
ཨྴ	0F15	TIBETAN LOGOTYPE SIGN CHAD RTAGS	NO	DISALLOWED
ཨྵ	0F16	TIBETAN LOGOTYPE SIGN LHAG RTAGS	NO	DISALLOWED
ཨྶ	0F17	TIBETAN ASTROLOGICAL SIGN SGRA GCAN -CHAR RTAGS	NO	DISALLOWED
ཨྷ	0F18	TIBETAN ASTROLOGICAL SIGN - KHYUD PA	NO	PVALID
ཨྸ	0F19	TIBETAN ASTROLOGICAL SIGN SDONG TSHUGS	NO	PVALID
ཨྐྵ	0F1A	TIBETAN SIGN RDEL DKAR GCIG	NO	DISALLOWED
ཨྺ	0F1B	TIBETAN SIGN RDEL DKAR GNYIS	NO	DISALLOWED
ཨྻ	0F1C	TIBETAN SIGN RDEL DKAR GSUM	NO	DISALLOWED
ཨྼ	0F1D	TIBETAN SIGN RDEL NAG GCIG	NO	DISALLOWED
ཨ྽	0F1E	TIBETAN SIGN RDEL NAG GNYIS	NO	DISALLOWED
ཨ྿	0F1F	TIBETAN SIGN RDEL DKAR RDEL NAG	NO	DISALLOWED
ཨྰ	0F20	TIBETAN DIGIT ZERO	YES	PVALID
ཨྱ	0F21	TIBETAN DIGIT ONE	YES	PVALID
ཨྲ	0F22	TIBETAN DIGIT TWO	YES	PVALID
ཨླ	0F23	TIBETAN DIGIT THREE	YES	PVALID
ཨྴ	0F24	TIBETAN DIGIT FOUR	YES	PVALID
ཨྵ	0F25	TIBETAN DIGIT FIVE	YES	PVALID
ཨྶ	0F26	TIBETAN DIGIT SIX	YES	PVALID
ཨྷ	0F27	TIBETAN DIGIT SEVEN	YES	PVALID
ཨྸ	0F28	TIBETAN DIGIT EIGHT	YES	PVALID
ཨྐྵ	0F29	TIBETAN DIGIT NINE	YES	PVALID

ཏ	0F2A	TIBETAN DIGIT HALF ONE	NO	DISALLOWED
ར	0F2B	TIBETAN DIGIT HALF TWO	NO	DISALLOWED
༣	0F2C	TIBETAN DIGIT HALF THREE	NO	DISALLOWED
པ	0F2D	TIBETAN DIGIT HALF FOUR	NO	DISALLOWED
ཕ	0F2E	TIBETAN DIGIT HALF FIVE	NO	DISALLOWED
བ	0F2F	TIBETAN DIGIT HALF SIX	NO	DISALLOWED
མ	0F30	TIBETAN DIGIT HALF SEVEN	NO	DISALLOWED
ཉ	0F31	TIBETAN DIGIT HALF EIGHT	NO	DISALLOWED
ཏ	0F32	TIBETAN DIGIT HALF NINE	NO	DISALLOWED
ཐ	0F33	TIBETAN DIGIT HALF ZERO	NO	DISALLOWED
ཨ	0F34	TIBETAN MARK BSDUS RTAGS	NO	DISALLOWED
ཨ ཨ	0F35	TIBETAN MARK NGAS BZUNG NYI ZLA	NO	PVALID
ཨ ཨ	0F36	TIBETAN MARK CARET -DZUD RTAGS BZHI MIG CAN	NO	DISALLOWED
ཨ ཨ	0F37	TIBETAN MARK NGAS BZUNG SGOR RTAGS	NO	PVALID
ཨ	0F38	TIBETAN MARK CHE MGO	NO	DISALLOWED
ཨ	0F39	TIBETAN MARK TSA -PHRU	YES	PVALID
ཨ	0F3A	TIBETAN MARK GUG RTAGS GYON	NO	DISALLOWED
ཨ	0F3B	TIBETAN MARK GUG RTAGS GYAS	NO	DISALLOWED
ཨ	0F3C	TIBETAN MARK ANG KHANG GYON	NO	DISALLOWED
ཨ	0F3D	TIBETAN MARK ANG KHANG GYAS	NO	DISALLOWED
ཨ	0F3E	TIBETAN SIGN YAR TSHES	NO	PVALID
ཨ	0F3F	TIBETAN SIGN MAR TSHES	NO	PVALID
ཀ	0F40	TIBETAN LETTER KA	YES	PVALID
ཁ	0F41	TIBETAN LETTER KHA	YES	PVALID
ག	0F42	TIBETAN LETTER GA	YES	PVALID













ཀ	0F43	TIBETAN LETTER GHA	YES	DISALLOWED
ང	0F44	TIBETAN LETTER NGA	YES	PVALID
ཅ	0F45	TIBETAN LETTER CA	YES	PVALID
ཆ	0F46	TIBETAN LETTER CHA	YES	PVALID
ཇ	0F47	TIBETAN LETTER JA	YES	PVALID
	0F48			UNASSIGNED
ཉ	0F49	TIBETAN LETTER NYA	YES	PVALID
ཊ	0F4A	TIBETAN LETTER TTA	YES	PVALID
ཋ	0F4B	TIBETAN LETTER TTHA	YES	PVALID
ཌ	0F4C	TIBETAN LETTER DDA	YES	PVALID
ཌྷ	0F4D	TIBETAN LETTER DDHA	YES	DISALLOWED
ཎ	0F4E	TIBETAN LETTER NNA	YES	PVALID
ཏ	0F4F	TIBETAN LETTER TA	YES	PVALID
ཐ	0F50	TIBETAN LETTER THA	YES	PVALID
ཏ	0F51	TIBETAN LETTER DA	YES	PVALID
ཐ	0F52	TIBETAN LETTER DHA	YES	DISALLOWED
ཏ	0F53	TIBETAN LETTER NA	YES	PVALID
ཏ	0F54	TIBETAN LETTER PA	YES	PVALID
ཏ	0F55	TIBETAN LETTER PHA	YES	PVALID
ཏ	0F56	TIBETAN LETTER BA	YES	PVALID
ཏ	0F57	TIBETAN LETTER BHA	YES	DISALLOWED
ཏ	0F58	TIBETAN LETTER MA	YES	PVALID
ཏ	0F59	TIBETAN LETTER TSA	YES	PVALID
ཏ	0F5A	TIBETAN LETTER TSHA	YES	PVALID
ཏ	0F5B	TIBETAN LETTER DZA	YES	PVALID
ཏ	0F5C	TIBETAN LETTER DZHA	YES	DISALLOWED

འ	0F5D	TIBETAN LETTER WA	YES	PVALID
ཅ	0F5E	TIBETAN LETTER ZHA	YES	PVALID
ཇ	0F5F	TIBETAN LETTER ZA	YES	PVALID
འ	0F60	TIBETAN LETTER -A	YES	PVALID
ཡ	0F61	TIBETAN LETTER YA	YES	PVALID
ཨ	0F62	TIBETAN LETTER RA	YES	PVALID
ལ	0F63	TIBETAN LETTER LA	YES	PVALID
ཤ	0F64	TIBETAN LETTER SHA	YES	PVALID
ས	0F65	TIBETAN LETTER SSA	YES	PVALID
མ	0F66	TIBETAN LETTER SA	YES	PVALID
ཨ	0F67	TIBETAN LETTER HA	YES	PVALID
།	0F68	TIBETAN LETTER A	YES	PVALID
ཤ	0F69	TIBETAN LETTER KSSA	YES	DISALLOWED
ཨ	0F6A	TIBETAN LETTER FIXED-FORM RA	YES	PVALID
	0F6B	<reserved>		UNASSIGNED
	0F6C	<reserved>		UNASSIGNED
	0F6D	<reserved>		UNASSIGNED
	0F6E	<reserved>		UNASSIGNED
	0F6F	<reserved>		UNASSIGNED
	0F70	<reserved>		UNASSIGNED
འ	0F71	TIBETAN VOWEL SIGN AA	YES	PVALID
འ	0F72	TIBETAN VOWEL SIGN I	YES	PVALID
འ	0F73	TIBETAN VOWEL SIGN II	YES	DISALLOWED
འ	0F74	TIBETAN VOWEL SIGN U	YES	PVALID
འ	0F75	TIBETAN VOWEL SIGN UU	YES	DISALLOWED
འ	0F76	TIBETAN VOWEL SIGN VOCALIC R	NO	DISALLOWED
འ	0F77	TIBETAN VOWEL SIGN VOCALIC RR	NO	DISALLOWED
འ	0F78	TIBETAN VOWEL SIGN VOCALIC L	NO	DISALLOWED

འཕྲུལ	0F79	TIBETAN VOWEL SIGN VOCALIC LL	NO	DISALLOWED
འཕྲུལ	0F7A	TIBETAN VOWEL SIGN E	YES	PVALID
འཕྲུལ	0F7B	TIBETAN VOWEL SIGN EE	YES	PVALID
འཕྲུལ	0F7C	TIBETAN VOWEL SIGN O	YES	PVALID
འཕྲུལ	0F7D	TIBETAN VOWEL SIGN OO	YES	PVALID
འཕྲུལ	0F7E	TIBETAN SIGN RJES SU NGA RO	NO	PVALID
འཕྲུལ	0F7F	TIBETAN SIGN RNAM BCAD	YES	PVALID
འཕྲུལ	0F80	TIBETAN VOWEL SIGN REVERSED I	YES	PVALID
འཕྲུལ	0F81	TIBETAN VOWEL SIGN REVERSED II	NO	DISALLOWED
འཕྲུལ	0F82	TIBETAN SIGN NYI ZLA NAA DA	NO	PVALID
འཕྲུལ	0F83	TIBETAN SIGN SNA LDAN	NO	PVALID
འཕྲུལ	0F84	TIBETAN MARK HALANTA	YES (but optional and will be removed)	PVALID
འཕྲུལ	0F85	TIBETAN MARK PALUTA	NO	DISALLOWED
འཕྲུལ	0F86	TIBETAN SIGN LCI RTAGS	NO	PVALID
འཕྲུལ	0F87	TIBETAN SIGN YANG RTAGS	NO	PVALID
འཕྲུལ	0F88	TIBETAN SIGN LCE TSA CAN	NO	PVALID
འཕྲུལ	0F89	TIBETAN SIGN MCHU CAN	NO	PVALID
འཕྲུལ	0F8A	TIBETAN SIGN GRU CAN RGYINGS	NO	PVALID
འཕྲུལ	0F8B	TIBETAN SIGN GRU MED RGYINGS	NO	PVALID
	0F8C	<reserved>		UNASSIGNED
	0F8D	<reserved>		UNASSIGNED
	0F8E	<reserved>		UNASSIGNED
	0F8F	<reserved>		UNASSIGNED
འཕྲུལ	0F90	TIBETAN SUBJOINED LETTER KA	YES	PVALID
འཕྲུལ	0F91	TIBETAN SUBJOINED LETTER KHA	YES	PVALID
འཕྲུལ	0F92	TIBETAN SUBJOINED LETTER GA	YES	PVALID

ཀམ	0F93	TIBETAN SUBJOINED LETTER GHA	YES	DISALLOWED
ཀཎ	0F94	TIBETAN SUBJOINED LETTER NGA	YES	PVALID
ཀཏ	0F95	TIBETAN SUBJOINED LETTER CA	YES	PVALID
ཀཐ	0F96	TIBETAN SUBJOINED LETTER CHA	NO	PVALID
ཀད	0F97	TIBETAN SUBJOINED LETTER JA	YES	PVALID
	0F98	<reserved>		UNASSIGNED
ཀར	0F99	TIBETAN SUBJOINED LETTER NYA	YES	PVALID
ཀལ	0F9A	TIBETAN SUBJOINED LETTER TTA	NO	PVALID
ཀཤ	0F9B	TIBETAN SUBJOINED LETTER TTHA	NO	PVALID
ཀཥ	0F9C	TIBETAN SUBJOINED LETTER DDA	NO	PVALID
ཀས	0F9D	TIBETAN SUBJOINED LETTER DDHA	NO	DISALLOWED
ཀཧ	0F9E	TIBETAN SUBJOINED LETTER NNA	NO	PVALID
ཀཨ	0F9F	TIBETAN SUBJOINED LETTER TA	YES	PVALID
ཀཀྵ	0FA0	TIBETAN SUBJOINED LETTER THA	YES	PVALID
ཀཪ	0FA1	TIBETAN SUBJOINED LETTER DA	YES	PVALID
ཀཫ	0FA2	TIBETAN SUBJOINED LETTER DHA	YES	DISALLOWED
ཀཬ	0FA3	TIBETAN SUBJOINED LETTER NA	YES	PVALID
ཀ཭	0FA4	TIBETAN SUBJOINED LETTER PA	YES	PVALID
ཀ཮	0FA5	TIBETAN SUBJOINED LETTER PHA	YES	PVALID
ཀཱ	0FA6	TIBETAN SUBJOINED LETTER BA	YES	PVALID
ཀི	0FA7	TIBETAN SUBJOINED LETTER BHA	YES	DISALLOWED
ཀཱི	0FA8	TIBETAN SUBJOINED LETTER MA	YES	PVALID
ཀུ	0FA9	TIBETAN SUBJOINED LETTER TSA	YES	PVALID
ཀཱུ	0FAA	TIBETAN SUBJOINED LETTER TSHA	YES	PVALID
ཀྲྀ	0FAB	TIBETAN SUBJOINED LETTER DZA	YES	PVALID
ཀཷ	0FAC	TIBETAN SUBJOINED LETTER DZHA	YES	DISALLOWED

འ	0FAD	TIBETAN SUBJOINED LETTER WA	YES	PVALID
ཅ	0FAE	TIBETAN SUBJOINED LETTER ZHA	YES	PVALID
ཎ	0FAF	TIBETAN SUBJOINED LETTER ZA	YES	PVALID
ཏ	0FB0	TIBETAN SUBJOINED LETTER -A	YES	PVALID
ཐ	0FB1	TIBETAN SUBJOINED LETTER YA	YES	PVALID
འ	0FB2	TIBETAN SUBJOINED LETTER RA	YES	PVALID
ཡ	0FB3	TIBETAN SUBJOINED LETTER LA	YES	PVALID
ལ	0FB4	TIBETAN SUBJOINED LETTER SHA	YES	PVALID
ཤ	0FB5	TIBETAN SUBJOINED LETTER SSA	YES	PVALID
ཇ	0FB6	TIBETAN SUBJOINED LETTER SA	YES	PVALID
ཉ	0FB7	TIBETAN SUBJOINED LETTER HA	YES	PVALID
ཏ	0FB8	TIBETAN SUBJOINED LETTER A	YES	PVALID
ཐ	0FB9	TIBETAN SUBJOINED LETTER KSSA	YES	DISALLOWED
འ	0FBA	TIBETAN SUBJOINED LETTER FIXED-FORM WA	YES	PVALID
ཡ	0FBB	TIBETAN SUBJOINED LETTER FIXED-FORM YA	YES	PVALID
ལ	0FBC	TIBETAN SUBJOINED LETTER FIXED-FORM RA	YES	PVALID
	0FBD	<reserved>		UNASSIGNED
མ	0FBE	TIBETAN KU RU KHA	NO	DISALLOWED
མ	0FBF	TIBETAN KU RU KHA BZHI MIG CAN	NO	DISALLOWED
འ	0FC0	TIBETAN CANTILLATION SIGN HEAVY BEAT	NO	DISALLOWED
འ	0FC1	TIBETAN CANTILLATION SIGN LIGHT BEAT	NO	DISALLOWED
འ	0FC2	TIBETAN CANTILLATION SIGN CANG TE-U	NO	DISALLOWED
འ	0FC3	TIBETAN CANTILLATION SIGN SBUB -CHAL	NO	DISALLOWED

	0FC4	TIBETAN SYMBOL DRIL BU	NO	DISALLOWED
	0FC5	TIBETAN SYMBOL RDO RJE	NO	DISALLOWED
	0FC6	TIBETAN SYMBOL PADMA GDAN	NO	PVALID
	0FC7	TIBETAN SYMBOL RDO RJE RGYA GRAM	NO	DISALLOWED
	0FC8	TIBETAN SYMBOL PHUR PA	NO	DISALLOWED
	0FC9	TIBETAN SYMBOL NOR BU	NO	DISALLOWED
	0FCA	TIBETAN SYMBOL NOR BU NYIS - KHYIL	NO	DISALLOWED
	0FCB	TIBETAN SYMBOL NOR BU GSUM - KHYIL	NO	DISALLOWED
	0FCC	TIBETAN SYMBOL NOR BU BZHI - KHYIL	NO	DISALLOWED
	0FCD	<reserved>		UNASSIGNED
	0FCE	<reserved>		UNASSIGNED
	0FCF	TIBETAN SIGN RDEL NAG GSUM	NO	DISALLOWED
	0FD0	TIBETAN MARK BSKA- SHOG GI MGO RGYAN	NO	DISALLOWED
	0FD1	TIBETAN MARK MNYAM YIG GI MGO RGYAN	NO	DISALLOWED

Appendix A.3: Khmer Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ក	1780	KHMER LETTER KA	YES	PVALID
ខ	1781	KHMER LETTER KHA	YES	PVALID
គ	1782	KHMER LETTER KO	YES	PVALID

ក	1783	KHMER LETTER KHO	YES	PVALID
ង	1784	KHMER LETTER NGO	YES	PVALID
ច	1785	KHMER LETTER CA	YES	PVALID
ឆ	1786	KHMER LETTER CHA	YES	PVALID
ជ	1787	KHMER LETTER CO	YES	PVALID
ឈ	1788	KHMER LETTER CHO	YES	PVALID
ញ	1789	KHMER LETTER NYO	YES	PVALID
ថ	1790	KHMER LETTER THA	YES	PVALID
ទ	1791	KHMER LETTER TO	YES	PVALID
ធ	1792	KHMER LETTER THO	YES	PVALID
ន	1793	KHMER LETTER NO	YES	PVALID
ប	1794	KHMER LETTER BA	YES	PVALID
ផ	1795	KHMER LETTER PHA	YES	PVALID
ព	1796	KHMER LETTER PO	YES	PVALID
ភ	1797	KHMER LETTER PHO	YES	PVALID
ម	1798	KHMER LETTER MO	YES	PVALID
យ	1799	KHMER LETTER YO	YES	PVALID
ដ	178A	KHMER LETTER DA	YES	PVALID
ឋ	178B	KHMER LETTER TTHA	YES	PVALID
ឌ	178C	KHMER LETTER DO	YES	PVALID
ឍ	178D	KHMER LETTER TTHO	YES	PVALID
ណ	178E	KHMER LETTER NNO	YES	PVALID
ត	178F	KHMER LETTER TA	YES	PVALID
រ	179A	KHMER LETTER RO	YES	PVALID
ល	179B	KHMER LETTER LO	YES	PVALID
វ	179C	KHMER LETTER VO	YES	PVALID
គ	179D	KHMER LETTER SHA	NO	PVALID

ថ	179E	KHMER LETTER SSO	NO	PVALID
ស	179F	KHMER LETTER SA	YES	PVALID
ហ	17A0	KHMER LETTER HA	YES	PVALID
ឡ	17A1	KHMER LETTER LA	YES	PVALID
អ	17A2	KHMER LETTER QA	YES	PVALID
អ	17A3	KHMER INDEPENDENT VOWEL QAQ	NO	PVALID
អ	17A4	KHMER INDEPENDENT VOWEL QAA	NO	PVALID
ឡ	17A5	KHMER INDEPENDENT VOWEL QI	YES	PVALID
ឡ	17A6	KHMER INDEPENDENT VOWEL QII	YES	PVALID
ឡ	17A7	KHMER INDEPENDENT VOWEL QU	YES	PVALID
ឡ	17A8	KHMER INDEPENDENT VOWEL QUK	YES	PVALID
ឡ	17A9	KHMER INDEPENDENT VOWEL QUU	YES	PVALID
ឡ	17AA	KHMER INDEPENDENT VOWEL QUUV	YES	PVALID
ឡ	17AB	KHMER INDEPENDENT VOWEL RY	YES	PVALID
ឡ	17AC	KHMER INDEPENDENT VOWEL RYY	YES	PVALID
ឡ	17AD	KHMER INDEPENDENT VOWEL LY	YES	PVALID
ឡ	17AE	KHMER INDEPENDENT VOWEL LYY	YES	PVALID
ឡ	17AF	KHMER INDEPENDENT VOWEL QE	YES	PVALID
ឡ	17B0	KHMER INDEPENDENT VOWEL QAI	YES	PVALID
ឡ	17B1	KHMER INDEPENDENT VOWEL QOO TYPE ONE	YES	PVALID
ឡ	17B2	KHMER INDEPENDENT VOWEL QOO TYPE TWO	YES	PVALID
ឡ	17B3	KHMER INDEPENDENT VOWEL QAU	YES	PVALID
៊	17B4	KHMER VOWEL INHERENT AQ	NO	CONTEXTO
៊	17B5	KHMER VOWEL INHERENT AA	NO	CONTEXTO
ា	17B6	KHMER VOWEL SIGN AA	YES	PVALID
ា	17B7	KHMER VOWEL SIGN I	YES	PVALID

◌ ◌	17B8	KHMER VOWEL SIGN II	YES	PVALID
◌ ◌	17B9	KHMER VOWEL SIGN Y	YES	PVALID
◌ ◌	17BA	KHMER VOWEL SIGN YY	YES	PVALID
◌ ◌	17BB	KHMER VOWEL SIGN U	YES	PVALID
◌ ◌	17BC	KHMER VOWEL SIGN UU	YES	PVALID
◌ ◌	17BD	KHMER VOWEL SIGN UA	YES	PVALID
◌ ◌	17BE	KHMER VOWEL SIGN OE	YES	PVALID
◌ ◌	17BF	KHMER VOWEL SIGN YA	YES	PVALID
◌ ◌	17C0	KHMER VOWEL SIGN IE	YES	PVALID
◌ ◌	17C1	KHMER VOWEL SIGN E	YES	PVALID
◌ ◌	17C2	KHMER VOWEL SIGN AE	YES	PVALID
◌ ◌	17C3	KHMER VOWEL SIGN AI	YES	PVALID
◌ ◌	17C4	KHMER VOWEL SIGN OO	YES	PVALID
◌ ◌	17C5	KHMER VOWEL SIGN AU	YES	PVALID
◌ ◌	17C6	KHMER SIGN NIKAHIT	YES	PVALID
◌ ◌	17C7	KHMER SIGN REAHMUK	YES	PVALID
◌ ◌	17C8	KHMER SIGN YUUKALEAPINTU	YES	PVALID
◌ ◌	17C9	KHMER SIGN MUUSIKATOAN	YES	PVALID
◌ ◌	17CA	KHMER SIGN TRIISAP	YES	PVALID
◌ ◌	17CB	KHMER SIGN BANTOC	YES	PVALID
◌ ◌	17CC	KHMER SIGN ROBAT	YES	PVALID
◌ ◌	17CD	KHMER SIGN TOANDAKHIAT	YES	PVALID
◌ ◌	17CE	KHMER SIGN KAKABAT	NO	PVALID
◌ ◌	17CF	KHMER SIGN AHSDA	YES	PVALID
◌ ◌	17D0	KHMER SIGN SAMYOK SANNYA	YES	PVALID
◌ ◌	17D1	KHMER SIGN VIRIAM	NO	PVALID
◌ ◌	17D2	KHMER SIGN COENG	YES	PVALID

០	17D3	KHMER SIGN BATHAMASAT	NO	PVALID
១	17D4	KHMER SIGN KHAN	NO	DISALLOWED
២	17D5	KHMER SIGN BARIYOOSAN	NO	DISALLOWED
៣	17D6	KHMER SIGN CAMNUC PII KUUH	NO	DISALLOWED
៤	17D7	KHMER SIGN LEK TOO	YES	PVALID
៥	17D8	KHMER SIGN BEYYAL	NO	DISALLOWED
៦	17D9	KHMER SIGN PHNAEK MUAN	NO	DISALLOWED
៧	17DA	KHMER SIGN KOOMUUT	NO	DISALLOWED
៨	17DB	KHMER CURRENCY SYMBOL RIEL	NO	DISALLOWED
៩	17DC	KHMER SIGN AVAKRAHASANYA	NO	PVALID
០	17DD	KHMER SIGN ATTHACAN	NO	PVALID
	17DE	<reserved>		UNASSIGNED
	17DF	<reserved>		UNASSIGNED
០	17E0	KHMER DIGIT ZERO	YES	PVALID
១	17E1	KHMER DIGIT ONE	YES	PVALID
២	17E2	KHMER DIGIT TWO	YES	PVALID
៣	17E3	KHMER DIGIT THREE	YES	PVALID
៤	17E4	KHMER DIGIT FOUR	YES	PVALID
៥	17E5	KHMER DIGIT FIVE	YES	PVALID
៦	17E6	KHMER DIGIT SIX	YES	PVALID
៧	17E7	KHMER DIGIT SEVEN	YES	PVALID
៨	17E8	KHMER DIGIT EIGHT	YES	PVALID
៩	17E9	KHMER DIGIT NINE	YES	PVALID
	17EA	<reserved>		UNASSIGNED
	17EB	<reserved>		UNASSIGNED
	17EC	<reserved>		UNASSIGNED
	17ED	<reserved>		UNASSIGNED
	17EE	<reserved>		UNASSIGNED

	17EF	<reserved>		UNASSIGNED
០	17F0	KHMER SYMBOL LEK ATTAK SON	NO	DISALLOWED
្ក	17F1	KHMER SYMBOL LEK ATTAK MUOY	NO	DISALLOWED
្ខ	17F2	KHMER SYMBOL LEK ATTAK PII	NO	DISALLOWED
្គ	17F3	KHMER SYMBOL LEK ATTAK BEI	NO	DISALLOWED
្ឃ	17F4	KHMER SYMBOL LEK ATTAK BUON	NO	DISALLOWED
្ង	17F5	KHMER SYMBOL LEK ATTAK PRAM	NO	DISALLOWED
្ច	17F6	KHMER SYMBOL LEK ATTAK PRAM-MUOY	NO	DISALLOWED
្ឆ	17F7	KHMER SYMBOL LEK ATTAK PRAM-PII	NO	DISALLOWED
្ជ	17F8	KHMER SYMBOL LEK ATTAK PRAM-BEI	NO	DISALLOWED
្ឈ	17F9	KHMER SYMBOL LEK ATTAK PRAM-BUON	NO	DISALLOWED
០ ១	19E0	KHMER SYMBOL PATHAMASAT	NO	DISALLOWED
១ ១	19E1	KHMER SYMBOL MUOY KOET	NO	DISALLOWED
២ ១	19E2	KHMER SYMBOL PII KOET	NO	DISALLOWED
៣ ១	19E3	KHMER SYMBOL BEI KOET	NO	DISALLOWED
៤ ១	19E4	KHMER SYMBOL BUON KOET	NO	DISALLOWED
៥ ១	19E5	KHMER SYMBOL PRAM KOET	NO	DISALLOWED
៦ ១	19E6	KHMER SYMBOL PRAM- MUOY KOET	NO	DISALLOWED
៧ ១	19E7	KHMER SYMBOL PRAM-PII KOET	NO	DISALLOWED
៨ ១	19E8	KHMER SYMBOL PRAM-BEI KOET	NO	DISALLOWED
៩ ១	19E9	KHMER SYMBOL PRAM- BUON KOET	NO	DISALLOWED
១០ ១	19EA	KHMER SYMBOL DAP KOET	NO	DISALLOWED
១១ ១	19EB	KHMER SYMBOL DAP- MUOY KOET	NO	DISALLOWED
១២ ១	19EC	KHMER SYMBOL DAP-PII KOET	NO	DISALLOWED
១៣ ១	19ED	KHMER SYMBOL DAP-BEI KOET	NO	DISALLOWED

១៤ ១	19EE	KHMER SYMBOL DAP- BUON KOET	NO	DISALLOWED
១៤ ១	19EF	KHMER SYMBOL DAP- PRAM KOET	NO	DISALLOWED
៤ ៤	19F0	KHMER SYMBOL TUTEYASAT	NO	DISALLOWED
១ ១	19F1	KHMER SYMBOL MUOY ROC	NO	DISALLOWED
១ ២	19F2	KHMER SYMBOL PII ROC	NO	DISALLOWED
១ ៣	19F3	KHMER SYMBOL BEI ROC	NO	DISALLOWED
១ ៤	19F4	KHMER SYMBOL BUON ROC	NO	DISALLOWED
១ ៥	19F5	KHMER SYMBOL PRAM ROC	NO	DISALLOWED
១ ៦	19F6	KHMER SYMBOL PRAM- MUOY ROC	NO	DISALLOWED
១ ៧	19F7	KHMER SYMBOL PRAM-PII ROC	NO	DISALLOWED
១ ៨	19F8	KHMER SYMBOL PRAM-BEI ROC	NO	DISALLOWED
១ ៩	19F9	KHMER SYMBOL PRAM- BUON ROC	NO	DISALLOWED
១ ១០	19FA	KHMER SYMBOL DAP ROC	NO	DISALLOWED
១ ១១	19FB	KHMER SYMBOL DAP- MUOY ROC	NO	DISALLOWED
១ ១២	19FC	KHMER SYMBOL DAP-PII ROC	NO	DISALLOWED
១ ១៣	19FD	KHMER SYMBOL DAP-BEI ROC	NO	DISALLOWED
១ ១៤	19FE	KHMER SYMBOL DAP- BUON ROC	NO	DISALLOWED
១ ១៥	19FF	KHMER SYMBOL DAP- PRAM ROC	NO	DISALLOWED

Appendix A.4: Lao Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ກ	0E81	LAO LETTER KO	YES	PVALID
ຂ	0E82	LAO LETTER KHO SUNG	YES	PVALID
	0E83	<reserved>		UNASSIGNED
ຄ	0E84	LAO LETTER KHO TAM	YES	PVALID
	0E85	<reserved>		UNASSIGNED
	0E86	<reserved>		UNASSIGNED
ງ	0E87	LAO LETTER NGO	YES	PVALID

๑	0E88	LAO LETTER CO	YES	PVALID
	0E89	<reserved>		UNASSIGNED
๒	0E8A	LAO LETTER SO TAM	YES	PVALID
	0E8B	<reserved>		UNASSIGNED
	0E8C	<reserved>		UNASSIGNED
๓	0E8D	LAO LETTER NYO	YES	PVALID
	0E8E	<reserved>		UNASSIGNED
	0E8F	<reserved>		UNASSIGNED
	0E90	<reserved>		UNASSIGNED
	0E91	<reserved>		UNASSIGNED
	0E92	<reserved>		UNASSIGNED
	0E93	<reserved>		UNASSIGNED
๔	0E94	LAO LETTER DO	YES	PVALID
๕	0E95	LAO LETTER TO	YES	PVALID
๖	0E96	LAO LETTER THO SUNG	YES	PVALID
๗	0E97	LAO LETTER THO TAM	YES	PVALID
	0E98	<reserved>		UNASSIGNED
๘	0E99	LAO LETTER DISALLOWED	YES	PVALID
๙	0E9A	LAO LETTER BO	YES	PVALID
๐	0E9B	LAO LETTER PO	YES	PVALID
๑	0E9C	LAO LETTER PHO SUNG	YES	PVALID
๒	0E9D	LAO LETTER FO TAM	YES	PVALID
๓	0E9E	LAO LETTER PHO TAM	YES	PVALID
๔	0E9F	LAO LETTER FO SUNG	YES	PVALID
	0EA0	<reserved>		UNASSIGNED
๕	0EA1	LAO LETTER MO	YES	PVALID
๖	0EA2	LAO LETTER YO	YES	PVALID
๗	0EA3	LAO LETTER LO LING	YES	PVALID
	0EA4	<reserved>		UNASSIGNED
๘	0EA5	LAO LETTER LO LOOT	YES	PVALID
	0EA6	<reserved>		UNASSIGNED
๙	0EA7	LAO LETTER WO	YES	PVALID
	0EA8	<reserved>		UNASSIGNED
	0EA9	<reserved>		UNASSIGNED
๐	0EAA	LAO LETTER SO SUNG	YES	PVALID
๑	0EAB	LAO LETTER HO SUNG	YES	PVALID
	0EAC	<reserved>		UNASSIGNED

ອ	0EAD	LAO LETTER O	YES	PVALID
ຮ	0EAE	LAO LETTER HO TAM	YES	PVALID
ຮ	0EAF	LAO ELLIPSIS	NO	PVALID
໊	0EB0	LAO VOWEL SIGN A	YES	PVALID
໋	0EB1	LAO VOWEL SIGN MAI KAN	YES	PVALID
໌	0EB2	LAO VOWEL SIGN AA	YES	PVALID
ໍ	0EB3	LAO VOWEL SIGN AM	YES	DISALLOWED
໎	0EB4	LAO VOWEL SIGN I	YES	PVALID
໏	0EB5	LAO VOWEL SIGN II	YES	PVALID
໐	0EB6	LAO VOWEL SIGN Y	YES	PVALID
໑	0EB7	LAO VOWEL SIGN YY	YES	PVALID
໒	0EB8	LAO VOWEL SIGN U	YES	PVALID
໓	0EB9	LAO VOWEL SIGN UU	YES	PVALID
	0EBA	<reserved>		UNASSIGNED
໔	0EBB	LAO VOWEL SIGN MAI KON	YES	PVALID
໕	0EBC	LAO SEMIVOWEL SIGN LO	YES	PVALID
໖	0EBD	LAO SEMI VOWEL SIGN NYO	YES	PVALID
	0EBE	<reserved>		UNASSIGNED
	0EBF	<reserved>		UNASSIGNED
໗	0EC0	LAO VOWEL SIGN	YES	PVALID
໘	0EC1	LAO VOWEL SIGN EI	YES	PVALID
໙	0EC2	LAO VOWEL SIGN O	YES	PVALID
໐	0EC3	LAO VOWEL SIGN AY	YES	PVALID
໑	0EC4	LAO VOWEL SIGN AI	YES	PVALID
	0EC5	<reserved>		UNASSIGNED
໒	0EC6	LAO KO LA	NO	PVALID
	0EC7	<reserved>		UNASSIGNED
໓	0EC8	LAO TONE MAI EK	YES	PVALID
໔	0EC9	LAO TONE MAI THO	YES	PVALID
໕	0ECA	LAO TONE TI	YES	PVALID
໖	0ECB	LAO TONE MAI CATAWA	YES	PVALID
໗	0ECC	LAO CANCELLATION MARK	YES	PVALID
໘	0ECD	LAO NIGGAHITA	YES	PVALID
	0ECE	<reserved>		UNASSIGNED
	0ECF	<reserved>		UNASSIGNED

໐	0ED0	LAO DIGIT ZERO	YES	PVALID
໑	0ED1	LAO DIGIT ONE	YES	PVALID
໒	0ED2	LAO DIGIT TWO	YES	PVALID
໓	0ED3	LAO DIGIT THREE	YES	PVALID
໔	0ED4	LAO DIGIT FOUR	YES	PVALID
໕	0ED5	LAO DIGIT FIVE	YES	PVALID
໖	0ED6	LAO DIGIT SIX	YES	PVALID
໗	0ED7	LAO DIGIT SEVEN	YES	PVALID
໘	0ED8	LAO DIGIT EIGHT	YES	PVALID
໙	0ED9	LAO DIGIT NINE	YES	PVALID
	0EDA	<reserved>		UNASSIGNED
	0EDB	<reserved>		UNASSIGNED
໙	0EDC	LAO LETTER HO NO	YES (variant of base character)	DISALLOWED
໙	0EDD	LAO LETTER HO MO	YES (variant of base character)	DISALLOWED

Appendix A.5: Mongolian Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
0	30	DIGIT ZERO	YES	PVALID
1	31	DIGIT ONE	YES	PVALID
2	32	DIGIT TWO	YES	PVALID
3	33	DIGIT THREE	YES	PVALID
4	34	DIGIT FOUR	YES	PVALID
5	35	DIGIT FIVE	YES	PVALID
6	36	DIGIT SIX	YES	PVALID
7	37	DIGIT SEVEN	YES	PVALID
8	38	DIGIT EIGHT	YES	PVALID
9	39	DIGIT NINE	YES	PVALID
Ё	400	CYRILLIC CAPITAL LETTER IE WITH GRAVE	NO	DISALLOWED
Ё	401	CYRILLIC CAPITAL LETTER IO	NO	DISALLOWED
Ђ	402	CYRILLIC CAPITAL LETTER DJE	NO	DISALLOWED
Ѓ	403	CYRILLIC CAPITAL LETTER GJE	NO	DISALLOWED
Є	404	CYRILLIC CAPITAL LETTER UKRAINIAN IE	NO	DISALLOWED

Ѕ	405	CYRILLIC CAPITAL LETTER DZE	NO	DISALLOWED
І	406	CYRILLIC CAPITAL LETTER BYELORUSSIAN-UKRAINIANI	NO	DISALLOWED
Ї	407	CYRILLIC CAPITAL LETTER YI	NO	DISALLOWED
Ј	408	CYRILLIC CAPITAL LETTER JE	NO	DISALLOWED
Љ	409	CYRILLIC CAPITAL LETTER LJE	NO	DISALLOWED
Њ	40A	CYRILLIC CAPITAL LETTER NJE	NO	DISALLOWED
Ћ	40B	CYRILLIC CAPITAL LETTER TSHE	NO	DISALLOWED
Ќ	40C	CYRILLIC CAPITAL LETTER KJE	NO	DISALLOWED
Ѳ	40D	CYRILLIC CAPITAL LETTER I WITH GRAVE	NO	DISALLOWED
ѳ	40E	CYRILLIC CAPITAL LETTER SHORT U	NO	DISALLOWED
Ѵ	40F	CYRILLIC CAPITAL LETTER DZHE	NO	DISALLOWED
А	410	CYRILLIC CAPITAL LETTER A	NO	DISALLOWED
Б	411	CYRILLIC CAPITAL LETTER BE	NO	DISALLOWED
В	412	CYRILLIC CAPITAL LETTER VE	NO	DISALLOWED
Г	413	CYRILLIC CAPITAL LETTER GHE	NO	DISALLOWED
Д	414	CYRILLIC CAPITAL LETTER DE	NO	DISALLOWED
Е	415	CYRILLIC CAPITAL LETTER IE	NO	DISALLOWED
Ж	416	CYRILLIC CAPITAL LETTER ZHE	NO	DISALLOWED
З	417	CYRILLIC CAPITAL LETTER ZE	NO	DISALLOWED
И	418	CYRILLIC CAPITAL LETTER I	NO	DISALLOWED
Ѳ	419	CYRILLIC CAPITAL LETTER SHORT I	NO	DISALLOWED
К	041A	CYRILLIC CAPITAL LETTER KA	NO	DISALLOWED
Л	041B	CYRILLIC CAPITAL LETTER EL	NO	DISALLOWED
М	041C	CYRILLIC CAPITAL LETTER EM	NO	DISALLOWED
Н	041D	CYRILLIC CAPITAL LETTER EN	NO	DISALLOWED

О	041E	CYRILLIC CAPITAL LETTER O	NO	DISALLOWED
П	041F	CYRILLIC CAPITAL LETTER PE	NO	DISALLOWED
Р	420	CYRILLIC CAPITAL LETTER ER	NO	DISALLOWED
С	421	CYRILLIC CAPITAL LETTER ES	NO	DISALLOWED
Т	422	CYRILLIC CAPITAL LETTER TE	NO	DISALLOWED
У	423	CYRILLIC CAPITAL LETTER U	NO	DISALLOWED
Ф	424	CYRILLIC CAPITAL LETTER EF	NO	DISALLOWED
Х	425	CYRILLIC CAPITAL LETTER HA	NO	DISALLOWED
Ц	426	CYRILLIC CAPITAL LETTER TSE	NO	DISALLOWED
Ч	427	CYRILLIC CAPITAL LETTER CHE	NO	DISALLOWED
Ш	428	CYRILLIC CAPITAL LETTER SHA	NO	DISALLOWED
Щ	429	CYRILLIC CAPITAL LETTER SHCHA	NO	DISALLOWED
Ъ	042A	CYRILLIC CAPITAL LETTER HARD SIGN	NO	DISALLOWED
Ы	042B	CYRILLIC CAPITAL LETTER YERU	NO	DISALLOWED
Ь	042C	CYRILLIC CAPITAL LETTER SOFT SIGN	NO	DISALLOWED
Э	042D	CYRILLIC CAPITAL LETTER E	NO	DISALLOWED
Ю	042E	CYRILLIC CAPITAL LETTER YU	NO	DISALLOWED
Я	042F	CYRILLIC CAPITAL LETTER YA	NO	DISALLOWED
а	430	CYRILLIC SMALL LETTER A	YES	PVALID
б	431	CYRILLIC SMALL LETTER BE	YES	PVALID
в	432	CYRILLIC SMALL LETTER VE	YES	PVALID
г	433	CYRILLIC SMALL LETTER GHE	YES	PVALID
д	434	CYRILLIC SMALL LETTER DE	YES	PVALID
е	435	CYRILLIC SMALL LETTER IE	YES	PVALID
ж	436	CYRILLIC SMALL LETTER ZHE	YES	PVALID
з	437	CYRILLIC SMALL LETTER ZE	YES	PVALID

и	438	CYRILLIC SMALL LETTER I	YES	PVALID
й	439	CYRILLIC SMALL LETTER SHORT I	YES	PVALID
к	043A	CYRILLIC SMALL LETTER KA	YES	PVALID
л	043B	CYRILLIC SMALL LETTER EL	YES	PVALID
м	043C	CYRILLIC SMALL LETTER EM	YES	PVALID
н	043D	CYRILLIC SMALL LETTER EN	YES	PVALID
о	043E	CYRILLIC SMALL LETTER O	YES	PVALID
п	043F	CYRILLIC SMALL LETTER PE	YES	PVALID
р	440	CYRILLIC SMALL LETTER ER	YES	PVALID
с	441	CYRILLIC SMALL LETTER ES	YES	PVALID
т	442	CYRILLIC SMALL LETTER TE	YES	PVALID
у	443	CYRILLIC SMALL LETTER U	YES	PVALID
ф	444	CYRILLIC SMALL LETTER EF	YES	PVALID
х	445	CYRILLIC SMALL LETTER HA	YES	PVALID
ц	446	CYRILLIC SMALL LETTER TSE	YES	PVALID
ч	447	CYRILLIC SMALL LETTER CHE	YES	PVALID
ш	448	CYRILLIC SMALL LETTER SHA	YES	PVALID
щ	449	CYRILLIC SMALL LETTER SHCHA	YES	PVALID
ъ	044A	CYRILLIC SMALL LETTER HARD SIGN	YES	PVALID
ы	044B	CYRILLIC SMALL LETTER YERU	YES	PVALID
ь	044C	CYRILLIC SMALL LETTER SOFT SIGN	YES	PVALID
э	044D	CYRILLIC SMALL LETTER E	YES	PVALID
ю	044E	CYRILLIC SMALL LETTER YU	YES	PVALID
я	044F	CYRILLIC SMALL LETTER YA	YES	PVALID
è	450	CYRILLIC SMALL LETTER IE WITH GRAVE	NO	PVALID

ё	451	CYRILLIC SMALL LETTER IO	YES	PVALID
ђ	452	CYRILLIC SMALL LETTER DJE	NO	PVALID
ѓ	453	CYRILLIC SMALL LETTER GJE	NO	PVALID
є	454	CYRILLIC SMALL LETTER UKRAINIAN IE	NO	PVALID
ѕ	455	CYRILLIC SMALL LETTER DZE	NO	PVALID
і	456	CYRILLIC SMALL LETTER BYELORUSSIAN-UKRAINIANI	NO	PVALID
ї	457	CYRILLIC SMALL LETTER YI	NO	PVALID
ј	458	CYRILLIC SMALL LETTER JE	NO	PVALID
љ	459	CYRILLIC SMALL LETTER LJE	NO	PVALID
њ	45A	CYRILLIC SMALL LETTER NJE	NO	PVALID
ћ	45B	CYRILLIC SMALL LETTER TSHE	NO	PVALID
ќ	45C	CYRILLIC SMALL LETTER KJE	NO	PVALID
ѝ	45D	CYRILLIC SMALL LETTER I WITH GRAVE	NO	PVALID
џ	45E	CYRILLIC SMALL LETTER SHORT U	NO	PVALID
ѡ	45F	CYRILLIC SMALL LETTER DZHE	NO	PVALID
Ѣ	460	CYRILLIC CAPITAL LETTER OMEGA	NO	DISALLOWED
ѣ	461	CYRILLIC SMALL LETTER OMEGA	NO	PVALID
Ѥ	462	CYRILLIC CAPITAL LETTER YAT	NO	DISALLOWED
ѥ	463	CYRILLIC SMALL LETTER YAT	NO	PVALID
Ѧ	464	CYRILLIC CAPITAL LETTER IOTIFIED E	NO	DISALLOWED
ѧ	465	CYRILLIC SMALL LETTER IOTIFIED E	NO	PVALID
Ѩ	466	CYRILLIC CAPITAL LETTER LITTLE YUS	NO	DISALLOWED
ѩ	467	CYRILLIC SMALL LETTER LITTLE YUS	NO	PVALID
Ѫ	468	CYRILLIC CAPITAL LETTER IOTIFIED	NO	DISALLOWED

		LITTLE YUS		
Ѧ	469	CYRILLIC SMALL LETTER IOTIFIED LITTLE YUS	NO	PVALID
Ѧ	046A	CYRILLIC CAPITAL LETTER BIG YUS	NO	DISALLOWED
Ѧ	046B	CYRILLIC SMALL LETTER BIG YUS	NO	PVALID
Ѧ	046C	CYRILLIC CAPITAL LETTER IOTIFIED BIG YUS	NO	DISALLOWED
Ѧ	046D	CYRILLIC SMALL LETTER IOTIFIED BIG YUS	NO	PVALID
Ѣ	046E	CYRILLIC CAPITAL LETTER KSI	NO	DISALLOWED
ѣ	046F	CYRILLIC SMALL LETTER KSI	NO	PVALID
Ѱ	470	CYRILLIC CAPITAL LETTER PSI	NO	DISALLOWED
ѱ	471	CYRILLIC SMALL LETTER PSI	NO	PVALID
Ѳ	472	CYRILLIC CAPITAL LETTER FITA	NO	DISALLOWED
ѳ	473	CYRILLIC SMALL LETTER FITA	NO	PVALID
Ѵ	474	CYRILLIC CAPITAL LETTER IZHITSA	NO	DISALLOWED
ѵ	475	CYRILLIC SMALL LETTER IZHITSA	NO	PVALID
Ѷ	476	CYRILLIC CAPITAL LETTER IZHITSA WITH DOUBLE GRAVE ACCENT	NO	DISALLOWED
ѷ	477	CYRILLIC SMALL LETTER IZHITSA WITH DOUBLE GRAVE ACCENT	NO	PVALID
Ѹ	478	CYRILLIC CAPITAL LETTER UK	NO	DISALLOWED
ѹ	479	CYRILLIC SMALL LETTER UK	NO	PVALID
Ѻ	047A	CYRILLIC CAPITAL LETTER ROUND OMEGA	NO	DISALLOWED
ѻ	047B	CYRILLIC SMALL LETTER ROUND OMEGA	NO	PVALID
Ѽ	047C	CYRILLIC CAPITAL LETTER OMEGA WITH TITLO	NO	DISALLOWED
ѽ	047D	CYRILLIC SMALL LETTER OMEGA WITH TITLO	NO	PVALID

Ѡ	047E	CYRILLIC CAPITAL LETTER OT	NO	DISALLOWED
ѡ	047F	CYRILLIC SMALL LETTER OT	NO	PVALID
Ѣ	480	CYRILLIC CAPITAL LETTER KOPPA	NO	DISALLOWED
ѣ	481	CYRILLIC SMALL LETTER KOPPA	NO	PVALID
Ѥ	482	CYRILLIC THOUSANDS SIGN	NO	DISALLOWED
ѥ	483	COMBINING CYRILLIC TITLO	NO	CONTEXTO
Ѧ	484	COMBINING CYRILLIC PALATALIZATION	NO	PVALID
ѧ	485	COMBINING CYRILLIC DASIA PNEUMATA	NO	PVALID
Ѩ	486	COMBINING CYRILLIC PSILI PNEUMATA	NO	PVALID
	487	<reserved>		UNASSIGNED
ѩ	488	COMBINING CYRILLIC HUNDRED THOUSANDS SIGN	NO	DISALLOWED
Ѫ	489	COMBINING CYRILLIC MILLIONS SIGN	NO	DISALLOWED
ѫ	048A	CYRILLIC CAPITAL LETTER SHORT I WITH TAIL	NO	DISALLOWED
Ѽ	048B	CYRILLIC SMALL LETTER SHORT I WITH TAIL	NO	PVALID
ѽ	048C	CYRILLIC CAPITAL LETTER SEMISOFT SIGN	NO	DISALLOWED
Ѿ	048D	CYRILLIC SMALL LETTER SEMISOFT SIGN	NO	PVALID
ѿ	048E	CYRILLIC CAPITAL LETTER ER WITH TICK	NO	DISALLOWED
Ѡ	048F	CYRILLIC SMALL LETTER ER WITH TICK	NO	PVALID
Ґ	490	CYRILLIC CAPITAL LETTER GHE WITH UPTURN (present in WGL4)	NO	DISALLOWED
ґ	491	CYRILLIC SMALL LETTER GHE WITH UPTURN (present in	NO	PVALID

		WGL4)		
Ғ	492	CYRILLIC CAPITAL LETTER GHE WITH STROKE	NO	DISALLOWED
ғ	493	CYRILLIC SMALL LETTER GHE WITH STROKE	NO	PVALID
Ғ	494	CYRILLIC CAPITAL LETTER GHE WITH MIDDLE HOOK	NO	DISALLOWED
ғ	495	CYRILLIC SMALL LETTER GHE WITH MIDDLE HOOK	NO	PVALID
Ҙ	496	CYRILLIC CAPITAL LETTER ZHE WITH DESCENDER	NO	DISALLOWED
ҙ	497	CYRILLIC SMALL LETTER ZHE WITH DESCENDER	NO	PVALID
Ҙ	498	CYRILLIC CAPITAL LETTER ZE WITH DESCENDER	NO	DISALLOWED
ҙ	499	CYRILLIC SMALL LETTER ZE WITH DESCENDER	NO	PVALID
Ҙ	049A	CYRILLIC CAPITAL LETTER KA WITH DESCENDER	NO	DISALLOWED
ҙ	049B	CYRILLIC SMALL LETTER KA WITH DESCENDER	NO	PVALID
Ҙ	049C	CYRILLIC CAPITAL LETTER KA WITH VERTICAL STROKE	NO	DISALLOWED
ҙ	049D	CYRILLIC SMALL LETTER KA WITH VERTICAL STROKE	NO	PVALID
Ҙ	049E	CYRILLIC CAPITAL LETTER KA WITH STROKE	NO	DISALLOWED
ҙ	049F	CYRILLIC SMALL LETTER KA WITH STROKE	NO	PVALID
Ҙ	04A0	CYRILLIC CAPITAL LETTER BASHKIR KA	NO	DISALLOWED
ҙ	04A1	CYRILLIC SMALL LETTER BASHKIR KA	NO	PVALID
Ҙ	04A2	CYRILLIC CAPITAL LETTER EN WITH DESCENDER	NO	DISALLOWED
ҙ	04A3	CYRILLIC SMALL LETTER EN WITH	NO	PVALID

		DESCENDER		
Ғ	04A4	CYRILLIC CAPITAL LIGATURE EN GHE	NO	DISALLOWED
ғ	04A5	CYRILLIC SMALL LIGATURE EN GHE	NO	PVALID
Ғб	04A6	CYRILLIC CAPITAL LETTER PE WITH MIDDLE HOOK	NO	DISALLOWED
ғб	04A7	CYRILLIC SMALL LETTER PE WITH MIDDLE HOOK	NO	PVALID
Ӓ	04A8	CYRILLIC CAPITAL LETTER ABKHASIAN HA	NO	DISALLOWED
ӓ	04A9	CYRILLIC SMALL LETTER ABKHASIAN HA	NO	PVALID
Ҙ	04AA	CYRILLIC CAPITAL LETTER ES WITH DESCENDER	NO	DISALLOWED
ҙ	04AB	CYRILLIC SMALL LETTER ES WITH DESCENDER	NO	PVALID
Ҙ	04AC	CYRILLIC CAPITAL LETTER TE WITH DESCENDER	NO	DISALLOWED
ҙ	04AD	CYRILLIC SMALL LETTER TE WITH DESCENDER	NO	PVALID
Ү	04AE	CYRILLIC CAPITAL LETTER STRAIGHT U	NO	DISALLOWED
ү	04AF	CYRILLIC SMALL LETTER STRAIGHT U	NO	PVALID
Ү	04B0	CYRILLIC CAPITAL LETTER STRAIGHT U WITH STROKE	NO	DISALLOWED
ү	04B1	CYRILLIC SMALL LETTER STRAIGHT U WITH STROKE	NO	PVALID
Ҳ	04B2	CYRILLIC CAPITAL LETTER HA WITH DESCENDER	NO	DISALLOWED
ҳ	04B3	CYRILLIC SMALL LETTER HA WITH DESCENDER	NO	PVALID
Ҙ	04B4	CYRILLIC CAPITAL LIGATURE TE TSE	NO	DISALLOWED
ҙ	04B5	CYRILLIC SMALL LIGATURE TE TSE	NO	PVALID
Ү	04B6	CYRILLIC CAPITAL LETTER CHE WITH DESCENDER	NO	DISALLOWED

ч	04B7	CYRILLIC SMALL LETTER CHE WITH DESCENDER	NO	PVALID
Ч	04B8	CYRILLIC CAPITAL LETTER CHE WITH VERTICAL STROKE	NO	DISALLOWED
ч	04B9	CYRILLIC SMALL LETTER CHE WITH VERTICAL STROKE	NO	PVALID
Һ	04BA	CYRILLIC CAPITAL LETTER SHHA	NO	DISALLOWED
һ	04BB	CYRILLIC SMALL LETTER SHHA	NO	PVALID
Ҁ	04BC	CYRILLIC CAPITAL LETTER ABKHASIAN CHE	NO	DISALLOWED
ҁ	04BD	CYRILLIC SMALL LETTER ABKHASIAN CHE	NO	PVALID
҂	04BE	CYRILLIC CAPITAL LETTER ABKHASIAN CHE WITH DESCENDER	NO	DISALLOWED
҃	04BF	CYRILLIC SMALL LETTER ABKHASIAN CHE WITH DESCENDER	NO	PVALID
І	04C0	CYRILLIC LETTER PALOCHKA	NO	DISALLOWED
Ж	04C1	CYRILLIC CAPITAL LETTER ZHE WITH BREVE	NO	DISALLOWED
ж	04C2	CYRILLIC SMALL LETTER ZHE WITH BREVE	NO	PVALID
Ҁ	04C3	CYRILLIC CAPITAL LETTER KA WITH HOOK	NO	DISALLOWED
ҁ	04C4	CYRILLIC SMALL LETTER KA WITH HOOK	NO	PVALID
҂	04C5	CYRILLIC CAPITAL LETTER EL WITH TAIL	NO	DISALLOWED
҃	04C6	CYRILLIC SMALL LETTER EL WITH TAIL	NO	PVALID
Ҁ	04C7	CYRILLIC CAPITAL LETTER EN WITH HOOK	NO	DISALLOWED
ҁ	04C8	CYRILLIC SMALL LETTER EN WITH HOOK	NO	PVALID
҂	04C9	CYRILLIC CAPITAL	NO	DISALLOWED

		LETTER EN WITH TAIL		
Ң	04CA	CYRILLIC SMALL LETTER EN WITH TAIL	NO	PVALID
Ҥ	04CB	CYRILLIC CAPITAL LETTER KHAKASSIAN CHE	NO	DISALLOWED
ҥ	04CC	CYRILLIC SMALL LETTER KHAKASSIAN CHE	NO	PVALID
Ҫ	04CD	CYRILLIC CAPITAL LETTER EM WITH TAIL	NO	DISALLOWED
Ҭ	04CE	CYRILLIC SMALL LETTER EM WITH TAIL	NO	PVALID
І	04CF	CYRILLIC SMALL LETTER PALOCHKA	NO	PVALID
Ӑ	04D0	CYRILLIC CAPITAL LETTER A WITH BREVE	NO	DISALLOWED
ӑ	04D1	CYRILLIC SMALL LETTER A WITH BREVE	NO	PVALID
Ӓ	04D2	CYRILLIC CAPITAL LETTER A WITH DIAERESIS	NO	DISALLOWED
ӓ	04D3	CYRILLIC SMALL LETTER A WITH DIAERESIS	NO	PVALID
Ӕ	04D4	CYRILLIC CAPITAL LIGATURE A IE	NO	DISALLOWED
ӕ	04D5	CYRILLIC SMALL LIGATURE A IE	NO	PVALID
Ӗ	04D6	CYRILLIC CAPITAL LETTER IE WITH BREVE	NO	DISALLOWED
ӗ	04D7	CYRILLIC SMALL LETTER IE WITH BREVE	NO	PVALID
Ә	04D8	CYRILLIC CAPITAL LETTER SCHWA	NO	DISALLOWED
ә	04D9	CYRILLIC SMALL LETTER SCHWA	NO	PVALID
Ӛ	04DA	CYRILLIC CAPITAL LETTER SCHWA WITH DIAERESIS	NO	DISALLOWED
ӛ	04DB	CYRILLIC SMALL LETTER SCHWA WITH DIAERESIS	NO	PVALID
Ә	04DC	CYRILLIC CAPITAL LETTER ZHE WITH	NO	DISALLOWED

		DIAERESIS		
Ӗ	04DD	CYRILLIC SMALL LETTER ZHE WITH DIAERESIS	NO	PVALID
Ӑ	04DE	CYRILLIC CAPITAL LETTER ZE WITH DIAERESIS	NO	DISALLOWED
ӓ	04DF	CYRILLIC SMALL LETTER ZE WITH DIAERESIS	NO	PVALID
Ӓ	4E0	CYRILLIC CAPITAL LETTER ABKHASIAN DZE	NO	DISALLOWED
ӑ	4E1	CYRILLIC SMALL LETTER ABKHASIAN DZE	NO	PVALID
Ӗ̄	4E2	CYRILLIC CAPITAL LETTER I WITH MACRON	NO	DISALLOWED
ӗ̄	4E3	CYRILLIC SMALL LETTER I WITH MACRON	NO	PVALID
Ӗ̆	4E4	CYRILLIC CAPITAL LETTER I WITH DIAERESIS	NO	DISALLOWED
ӗ̆	4E5	CYRILLIC SMALL LETTER I WITH DIAERESIS	NO	PVALID
Ӗ̈	4E6	CYRILLIC CAPITAL LETTER O WITH DIAERESIS	NO	DISALLOWED
ӗ̈	4E7	CYRILLIC SMALL LETTER O WITH DIAERESIS	NO	PVALID
Ө	4E8	CYRILLIC CAPITAL LETTER BARRED O	YES	DISALLOWED
ө	4E9	CYRILLIC SMALL LETTER BARRED O	YES	PVALID
Ӗ̆̈	04EA	CYRILLIC CAPITAL LETTER BARRED O WITH DIAERESIS	NO	DISALLOWED
ӗ̆̈	04EB	CYRILLIC SMALL LETTER BARRED O WITH DIAERESIS	NO	PVALID
Ӗ̆̄	04EC	CYRILLIC CAPITAL LETTER E WITH DIAERESIS	NO	DISALLOWED
ӗ̆̄	04ED	CYRILLIC SMALL LETTER E WITH DIAERESIS	NO	PVALID
Ӗ̄̄	04EE	CYRILLIC CAPITAL LETTER U WITH MACRON	NO	DISALLOWED
ӗ̄̄	04EF	CYRILLIC SMALL	NO	PVALID

		LETTER U WITH MACRON		
Ѹ	04F0	CYRILLIC CAPITAL LETTER U WITH DIAERESIS	NO	DISALLOWED
ѹ	04F1	CYRILLIC SMALL LETTER U WITH DIAERESIS	NO	PVALID
Ѻ	04F2	CYRILLIC CAPITAL LETTER U WITH DOUBLE ACUTE	NO	DISALLOWED
ѻ	04F3	CYRILLIC SMALL LETTER U WITH DOUBLE ACUTE	NO	PVALID
Ѽ	04F4	CYRILLIC CAPITAL LETTER CHE WITH DIAERESIS	NO	DISALLOWED
ѽ	04F5	CYRILLIC SMALL LETTER CHE WITH DIAERESIS	NO	PVALID
Ѿ	04F6	CYRILLIC CAPITAL LETTER GHE WITH DESCENDER	NO	DISALLOWED
ѿ	04F7	CYRILLIC SMALL LETTER GHE WITH DESCENDER	NO	PVALID
Ѡ	04F8	CYRILLIC CAPITAL LETTER YERU WITH DIAERESIS	NO	DISALLOWED
ѡ	04F9	CYRILLIC SMALL LETTER YERU WITH DIAERESIS	NO	PVALID
Ѣ	04FA	CYRILLIC CAPITAL LETTER GHE WITH STROKE AND HOOK	NO	DISALLOWED
ѣ	04FB	CYRILLIC SMALL LETTER GHE WITH STROKE AND HOOK	NO	PVALID
Ѥ	04FC	CYRILLIC CAPITAL LETTER HA WITH HOOK	NO	DISALLOWED
ѥ	04FD	CYRILLIC SMALL LETTER HA WITH HOOK	NO	PVALID
Ѧ	04FE	CYRILLIC CAPITAL LETTER HA WITH STROKE	NO	DISALLOWED
ѧ	04FF	CYRILLIC SMALL LETTER HA WITH STROKE	NO	PVALID
Ѩ	500	CYRILLIC CAPITAL LETTER KOMI DE	NO	DISALLOWED

д	501	CYRILLIC SMALL LETTER KOMI DE	YES	PVALID
Д	502	CYRILLIC CAPITAL LETTER KOMI DJE	NO	DISALLOWED
д	503	CYRILLIC SMALL LETTER KOMI DJE	YES	PVALID
З	504	CYRILLIC CAPITAL LETTER KOMI ZJE	NO	DISALLOWED
з	505	CYRILLIC SMALL LETTER KOMI ZJE	YES	PVALID
З	506	CYRILLIC CAPITAL LETTER KOMI DZJE	NO	DISALLOWED
з	507	CYRILLIC SMALL LETTER KOMI DZJE	YES	PVALID
Љ	508	CYRILLIC CAPITAL LETTER KOMI LJE	NO	DISALLOWED
љ	509	CYRILLIC SMALL LETTER KOMI LJE	YES	PVALID
Њ	050A	CYRILLIC CAPITAL LETTER KOMI NJE	NO	DISALLOWED
њ	050B	CYRILLIC SMALL LETTER KOMI NJE	YES	PVALID
Ѓ	050C	CYRILLIC CAPITAL LETTER KOMI SJE	NO	DISALLOWED
ѓ	050D	CYRILLIC SMALL LETTER KOMI SJE	YES	PVALID
Т	050E	CYRILLIC CAPITAL LETTER KOMI TJE	NO	DISALLOWED
т	050F	CYRILLIC SMALL LETTER KOMI TJE	YES	PVALID
Э	510	CYRILLIC CAPITAL LETTER REVERSED ZE	NO	DISALLOWED
э	511	CYRILLIC SMALL LETTER REVERSED ZE	YES	PVALID
Ј	512	CYRILLIC CAPITAL LETTER EL WITH HOOK	NO	DISALLOWED
ј	513	CYRILLIC SMALL LETTER EL WITH HOOK	YES	PVALID

Appendix A.6: Nepali Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
ॐ	901	DEVANAGARI SIGN CANDRABINDU	YES	PVALID

ं	902	DEVANAGARI SIGN ANUSVARA	YES	PVALID
ः	903	DEVANAGARI SIGN VISARGA	YES	PVALID
ऐ	904	DEVANAGARI LETTER SHORT A	NO	PVALID
अ	905	DEVANAGARI LETTER A	YES	PVALID
आ	906	DEVANAGARI LETTER AA	YES	PVALID
इ	907	DEVANAGARI LETTER I	YES	PVALID
ई	908	DEVANAGARI LETTER II	YES	PVALID
उ	909	DEVANAGARI LETTER U	YES	PVALID
ऐ	910	DEVANAGARI LETTER AI	YES	PVALID
ऑ	911	DEVANAGARI LETTER CANDRA O	NO	PVALID
ओ	912	DEVANAGARI LETTER SHORT O	NO	PVALID
ओ	913	DEVANAGARI LETTER O	YES	PVALID
औ	914	DEVANAGARI LETTER AU	YES	PVALID
क	915	DEVANAGARI LETTER KA	YES	PVALID
ख	916	DEVANAGARI LETTER KHA	YES	PVALID
ग	917	DEVANAGARI LETTER GA	YES	PVALID
घ	918	DEVANAGARI LETTER GHA	YES	PVALID
ङ	919	DEVANAGARI LETTER NGA	YES	PVALID
ठ	920	DEVANAGARI LETTER TTHA	YES	PVALID
ड	921	DEVANAGARI LETTER DDA	YES	PVALID

ढ	922	DEVANAGARI LETTER DDHA	YES	PVALID
ण	923	DEVANAGARI LETTER NNA	YES	PVALID
त	924	DEVANAGARI LETTER TA	YES	PVALID
थ	925	DEVANAGARI LETTER THA	YES	PVALID
द	926	DEVANAGARI LETTER DA	YES	PVALID
ध	927	DEVANAGARI LETTER DHA	YES	PVALID
न	928	DEVANAGARI LETTER NA	YES	PVALID
ॢ	929	DEVANAGARI LETTER NNNA	YES (variant of base character)	PVALID
र	930	DEVANAGARI LETTER RA	YES	PVALID
ॠ	931	DEVANAGARI LETTER RRA	NO	PVALID
ल	932	DEVANAGARI LETTER LA	YES	PVALID
ळ	933	DEVANAGARI LETTER LLA	YES (variant of base character)	PVALID
ॡ	934	DEVANAGARI LETTER LLLA	NO	PVALID
व	935	DEVANAGARI LETTER VA	YES	PVALID
श	936	DEVANAGARI LETTER SHA	YES	PVALID
ष	937	DEVANAGARI LETTER SSA	YES	PVALID
स	938	DEVANAGARI LETTER SA	YES	PVALID
ह	939	DEVANAGARI LETTER HA	YES	PVALID
ी	940	DEVANAGARI VOWEL SIGN II	YES	PVALID

ॊ	941	DEVANAGARI VOWEL SIGN U	YES	PVALID
ो	942	DEVANAGARI VOWEL SIGN UU	YES	PVALID
ौ	943	DEVANAGARI VOWEL SIGN VOCALIC R	YES	PVALID
्	944	DEVANAGARI VOWEL SIGN VOCALIC RR	NO	PVALID
ॎ	945	DEVANAGARI VOWEL SIGN CANDRA E	NO	PVALID
ॏ	946	DEVANAGARI VOWEL SIGN SHORT E	NO	PVALID
ॐ	947	DEVANAGARI VOWEL SIGN E	YES	PVALID
॑	948	DEVANAGARI VOWEL SIGN AI	YES	PVALID
॒	949	DEVANAGARI VOWEL SIGN CANDRA O	NO	PVALID
ॐ	950	DEVANAGARI OM	YES	PVALID
॑	951	DEVANAGARI STRESS SIGN UDATTA	NO	PVALID
॑	952	DEVANAGARI STRESS SIGN ANUDATTA	NO	PVALID
॑	953	DEVANAGARI GRAVE ACCENT	NO	PVALID
॑	954	DEVANAGARI ACUTE ACCENT	NO	PVALID
	955	<reserved>		UNASSIGNED
	956	<reserved>		UNASSIGNED
	957	<reserved>		UNASSIGNED
क	958	DEVANAGARI LETTER QA	NO	DISALLOWED
ख	959	DEVANAGARI LETTER KHHA	NO	DISALLOWED
ऋ	960	DEVANAGARI LETTER VOCALIC RR	NO	PVALID
ॠ	961	DEVANAGARI LETTER VOCALIC LL	NO	PVALID

ॐ	962	DEVANAGARI VOWEL SIGN VOCALIC L	NO	PVALID
ॐ	963	DEVANAGARI VOWEL SIGN VOCALIC LL	NO	PVALID
क।	964	DEVANAGARI DANDA	NO	DISALLOWED
क॥	965	DEVANAGARI DOUBLE DANDA	NO	DISALLOWED
०	966	DEVANAGARI DIGIT ZERO	YES	PVALID
१	967	DEVANAGARI DIGIT ONE	YES	PVALID
२	968	DEVANAGARI DIGIT TWO	YES	PVALID
३	969	DEVANAGARI DIGIT THREE	YES	PVALID
॰	970	DEVANAGARI ABBREVIATION SIGN	NO	DISALLOWED
	971	<reserved>		UNASSIGNED
	972	<reserved>		UNASSIGNED
	973	<reserved>		UNASSIGNED
	974	<reserved>		UNASSIGNED
	975	<reserved>		UNASSIGNED
	976	<reserved>		UNASSIGNED
	977	<reserved>		UNASSIGNED
	978	<reserved>		UNASSIGNED
	979	<reserved>		UNASSIGNED
ऊ	090A	DEVANAGARI LETTER UU	YES	PVALID
ऋ	090B	DEVANAGARI LETTER VOCALIC R	YES	PVALID
ॠ	090C	DEVANAGARI LETTER VOCALIC L	NO	PVALID
ँ	090D	DEVANAGARI LETTER CANDRA E	NO	PVALID
ऐ	090E	DEVANAGARI LETTER SHORT E	NO	PVALID
ॢ	090F	DEVANAGARI LETTER E	YES	PVALID

च	091A	DEVANAGARI LETTER CA	YES	PVALID
छ	091B	DEVANAGARI LETTER CHA	YES	PVALID
ज	091C	DEVANAGARI LETTER JA	YES	PVALID
झ	091D	DEVANAGARI LETTER JHA	YES	PVALID
ञ	091E	DEVANAGARI LETTER NYA	YES	PVALID
ट	091F	DEVANAGARI LETTER TTA	YES	PVALID
प	092A	DEVANAGARI LETTER PA	YES	PVALID
फ	092B	DEVANAGARI LETTER PHA	YES	PVALID
ब	092C	DEVANAGARI LETTER BA	YES	PVALID
भ	092D	DEVANAGARI LETTER BHA	YES	PVALID
म	092E	DEVANAGARI LETTER MA	YES	PVALID
य	092F	DEVANAGARI LETTER YA	YES	PVALID
	093A	<reserved>		UNASSIGNED
	093B	<reserved>		UNASSIGNED
◌्	093C	DEVANAGARI SIGN NUKTA	NO	PVALID
◌ः	093D	DEVANAGARI SIGN AVAGRAHA	NO	PVALID
◌ा	093E	DEVANAGARI VOWEL SIGN AA	YES	PVALID
◌ि	093F	DEVANAGARI VOWEL SIGN I	YES	PVALID
◌ी	094A	DEVANAGARI VOWEL SIGN SHORT O	NO	PVALID
◌ो	094B	DEVANAGARI VOWEL SIGN O	YES	PVALID
◌ौ	094C	DEVANAGARI VOWEL SIGN AU	YES	PVALID
◌ं	094D	DEVANAGARI SIGN VIRAMA	YES	PVALID
	094E	<reserved>		UNASSIGNED
	094F	<reserved>		UNASSIGNED

ग	095A	DEVANAGARI LETTER GHHA	NO	DISALLOWED
ज	095B	DEVANAGARI LETTER ZA	NO	DISALLOWED
ड	095C	DEVANAGARI LETTER DDDHA	NO	DISALLOWED
ढ	095D	DEVANAGARI LETTER RHA	NO	DISALLOWED
फ	095E	DEVANAGARI LETTER FA	NO	DISALLOWED
य	095F	DEVANAGARI LETTER YYA	NO	DISALLOWED
४	096A	DEVANAGARI DIGIT FOUR	YES	PVALID
५	096B	DEVANAGARI DIGIT FIVE	YES	PVALID
६	096C	DEVANAGARI DIGIT SIX	YES	PVALID
७	096D	DEVANAGARI DIGIT SEVEN	YES	PVALID
८	096E	DEVANAGARI DIGIT EIGHT	YES	PVALID
९	096F	DEVANAGARI DIGIT NINE	YES	PVALID
	097A	<reserved>		UNASSIGNED
ग	97B	DEVANAGARI LETTER GGA	NO	PVALID
ज	97C	DEVANAGARI LETTER JJA	NO	PVALID
२	97D	DEVANAGARI LETTER GLOTAL STOP	NO	PVALID
ड	97E	DEVANAGARI LETTER DDDA	NO	PVALID
ढ	97F	DEVANAGARI LETTER BBA	NO	PVALID

Appendix A.7: Pashto Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
٠	600	ARABIC NUMBER SIGN	NO	CONTEXT0
١	601	ARABIC SIGN SANAH	NO	CONTEXT0
٢	601	ARABIC FOOTNOTE MARKER	NO	CONTEXT0
٣	603	ARABIC SIGN SAFHA	NO	CONTEXT0
	604	<reserved>	UNASSIGNED	UNASSIGNED
	605	<reserved>	UNASSIGNED	UNASSIGNED
	606	<reserved>	UNASSIGNED	UNASSIGNED
	607	<reserved>	UNASSIGNED	UNASSIGNED
	608	<reserved>	UNASSIGNED	UNASSIGNED
	609	<reserved>	UNASSIGNED	UNASSIGNED
	060A	<reserved>	UNASSIGNED	UNASSIGNED
٤	060B	AFGHANI SIGN	NO	DISALLOWED
٥	060C	ARABIC COMMA	NO	DISALLOWED
٦	060D	ARABIC DATE SEPARATOR	NO	DISALLOWED
٧	060E	ARABIC POETIC VERSE SIGN	NO	DISALLOWED
٨	060F	ARABIC SIGN MISRA	NO	DISALLOWED
٩	610	ARABIC SIGN SALLALLAHOU ALAYHE WASSALLAM	YES	PVALID
١٠	611	ARABIC SIGN ALAYHE ASSALLAM	YES	PVALID
١١	612	ARABIC SIGN RAHMATULLAH ALAYHE	YES	PVALID
١٢	613	ARABIC SIGN RADI ALLAHOU ANHU	YES	PVALID
١٣	614	ARABIC SIGN TAKHALLUS	NO	PVALID

□	615	ARABIC SMALL HIGH TAH	NO	PVALID
	616	<reserved>		UNASSIGNED
	617	<reserved>		UNASSIGNED
	618	<reserved>		UNASSIGNED
	619	<reserved>		UNASSIGNED
	061A	<reserved>		UNASSIGNED
؛	061B	ARABIC SEMICOLON	NO	DISALLOWED
		<reserved>		UNASSIGNED
		<reserved>		UNASSIGNED
⋯	061E	ARABIC TRIPLE DOT PUNCTUATION MARK	NO	DISALLOWED
؟	061F	ARABIC QUESTION MARK	NO	DISALLOWED
	620	<reserved>		UNASSIGNED
ء	621	ARABIC LETTER HAMZA	YES	PVALID
آ	622	ARABIC LETTER ALEF WITH MADDA ABOVE	YES	PVALID
أ	623	ARABIC LETTER ALEF WITH HAMZA ABOVE	YES	PVALID
ؤ	624	ARABIC LETTER WAW WITH HAMZA ABOVE	YES	PVALID
إ	625	ARABIC LETTER ALEF WITH HAMZA BELOW	YES (variant of base character)	PVALID
ئ	626	ARABIC LETTER YEH WITH HAMZA ABOVE	YES	PVALID
ا	627	ARABIC LETTER ALEF	YES	PVALID
ب	628	ARABIC LETTER BEH	YES	PVALID
ة	629	ARABIC LETTER TEH MARBUTA	YES	PVALID
ت	062A	ARABIC LETTER TEH	YES	PVALID
ث	062B	ARABIC LETTER THEH	YES	PVALID
ج	062C	ARABIC LETTER JEEM	YES	PVALID
ح	062D	ARABIC LETTER HAH	YES	PVALID
خ	062E	ARABIC LETTER KHAH	YES	PVALID
د	062F	ARABIC LETTER DAL	YES	PVALID
ذ	630	ARABIC LETTER THAL	YES	PVALID
ر	631	ARABIC LETTER REH	YES	PVALID
ز	632	ARABIC LETTER ZAIN	YES	PVALID
س	633	ARABIC LETTER SEEN	YES	PVALID

ا	654	ARABIC HAMZA ABOVE	YES	PVALID
آ	655	ARABIC HAMZA BELOW	YES	PVALID
آ	656	ARABIC SUBSCRIPT ALEF	YES	PVALID
آ	657	ARABIC INVERTED DAMMA	NO	PVALID
آ	658	ARABIC MARK NOON GHUNNA	YES	PVALID
آ	659	ARABIC ZWARAKAY	NO	PVALID
آ	065A	ARABIC VOWEL SIGN SMALL V ABOVE	NO	PVALID
آ	065B	ARABIC VOWEL SIGN INVERTED SMALL V ABOVE	NO	PVALID
آ	065C	ARABIC VOWEL SIGN DOT BELOW	NO	PVALID
آ	065D	ARABIC REVERSED DAMMA	NO	PVALID
آ	065E	ARABIC FATHA WITH TWO DOTS	NO	PVALID
	065F	<reserved>		UNASSIGNED
٠	660	ARABIC-INDIC DIGIT ZERO	YES	PVALID
١	661	ARABIC-INDIC DIGIT ONE	YES	PVALID
٢	662	ARABIC-INDIC DIGIT TWO	YES	PVALID
٣	663	ARABIC-INDIC DIGIT THREE	YES	PVALID
٤	664	ARABIC-INDIC DIGIT FOUR	YES	PVALID
٥	665	ARABIC-INDIC DIGIT FIVE	YES	PVALID
٦	666	ARABIC-INDIC DIGIT SIX	YES	PVALID
٧	667	ARABIC-INDIC DIGIT SEVEN	YES	PVALID
٨	668	ARABIC-INDIC DIGIT EIGHT	YES	PVALID
٩	669	ARABIC-INDIC DIGIT NINE	YES	PVALID
٪	066A	ARABIC PERCENT SIGN	NO	DISALLOWED
٫	066B	ARABIC DECIMAL SEPARATOR	NO	DISALLOWED
٬	066C	ARABIC THOUSANDS SEPARATOR	NO	DISALLOWED

*	066D	ARABIC FIVE POINTED STAR	NO	DISALLOWED
ﺀ	066E	ARABIC LETTER DOTLESS BEH	NO	PVALID
ﻮ	066F	ARABIC LETTER DOTLESS QAF	NO	PVALID
ﺀ	670	ARABIC LETTER SUPERSCRIPIT ALEF	YES	PVALID
ﺀ	671	ARABIC LETTER ALEF WASLA	NO	PVALID
ﺀ	672	ARABIC LETTER ALEF WITH WAVY HAMZA ABOVE	NO	PVALID
ﺀ	673	ARABIC LETTER ALEF WITH WAVY HAMZA BELOW	NO	PVALID
ﺀ	674	ARABIC LETTER HIGH HAMZA	NO	PVALID
ﺀ	675	ARABIC LETTER HIGH HAMZA ALEF	NO	DISALLOWED
ﺀ	676	ARABIC LETTER HIGH HAMZA WAW	NO	DISALLOWED
ﺀ	677	ARABIC LETTER U WITH HAMZA ABOVE	NO	DISALLOWED
ﺀ	678	ARABIC LETTER HIGH HAMZA YEH	NO	DISALLOWED
ﺀ	679	ARABIC LETTER TTEH	YES (variant of base character)	PVALID
ﺀ	067A	ARABIC LETTER TTEHEH	NO	PVALID
ﺀ	067B	ARABIC LETTER BEEH	NO	PVALID
ﺀ	067C	ARABIC LETTER TEH WITH RING	YES	PVALID
ﺀ	067D	ARABIC LETTER TEH WITH THREE DOTS ABOVE DOWN	NO	PVALID
ﺀ	067E	ARABIC LETTER PEH	YES	PVALID
ﺀ	067F	ARABIC LETTER TEHEH	NO	PVALID
ﺀ	680	ARABIC LETTER BEHEH	NO	PVALID
ﺀ	681	ARABIC LETTER HAH WITH HAMZA ABOVE	NO	PVALID
ﺀ	682	ARABIC LETTER HAH WITH TWO DOTS VERTICAL ABO	NO	PVALID
ﺀ	683	ARABIC LETTER NYEH	NO	PVALID
ﺀ	684	ARABIC LETTER DYEH	NO	PVALID
ﺀ	685	ARABIC LETTER HAH WITH THREE DOTS ABOVE	YES	PVALID
ﺀ	686	ARABIC LETTER TCHEH	YES	PVALID
ﺀ	687	ARABIC LETTER TCHEHEH	NO	PVALID
ﺀ	688	ARABIC LETTER DDAL	YES (variant of base character)	PVALID
ﺀ	689	ARABIC LETTER DAL WITH RING	YES	PVALID
ﺀ	068A	ARABIC LETTER DAL WITH DOT BELOW	NO	PVALID

ڙ	068B	ARABIC LETTER DAL WITH DOT BELOW AND SMALL T	NO	PVALID
ڙ	068C	ARABIC LETTER DAHAL	NO	PVALID
ڙ	068D	ARABIC LETTER DDAHAL	NO	PVALID
ڙ	068E	ARABIC LETTER DUL	NO	PVALID
ڙ	068F	ARABIC LETTER DAL WITH THREE DOTS ABOVE DOWN	NO	PVALID
ڙ	690	ARABIC LETTER DAL WITH FOUR DOTS ABOVE	NO	PVALID
ڙ	691	ARABIC LETTER RREH	YES (variant of base character)	PVALID
ڙ	692	ARABIC LETTER REH WITH SMALL V	NO	PVALID
ڙ	693	ARABIC LETTER REH WITH RING	YES	PVALID
ڙ	694	ARABIC LETTER REH WITH DOT BELOW	NO	PVALID
ڙ	695	ARABIC LETTER REH WITH SMALL V BELOW	NO	PVALID
ڙ	696	ARABIC LETTER REH WITH DOT BELOW AND DOT ABO	YES	PVALID
ڙ	697	ARABIC LETTER REH WITH TWO DOTS ABOVE	NO	PVALID
ڙ	698	ARABIC LETTER JEH	YES	PVALID
ڙ	699	ARABIC LETTER REH WITH FOUR DOTS ABOVE	NO	PVALID
ڙ	069A	ARABIC LETTER SEEN WITH DOT BELOW AND DOT AB	YES	PVALID
ڙ	069B	ARABIC LETTER SEEN WITH THREE DOTS BELOW	NO	PVALID
ڙ	069C	ARABIC LETTER SEEN WITH THREE DOTS BELOW AND	NO	PVALID
ڙ	069D	ARABIC LETTER SAD WITH TWO DOTS BELOW	NO	PVALID
ڙ	069E	ARABIC LETTER SAD WITH THREE DOTS ABOVE	NO	PVALID
ڙ	069F	ARABIC LETTER TAH WITH THREE DOTS ABOVE	NO	PVALID
ڙ	06A0	ARABIC LETTER AIN WITH THREE DOTS ABOVE	NO	PVALID
ڙ	06A1	ARABIC LETTER DOTLESS FEH	NO	PVALID
ڙ	06A2	ARABIC LETTER FEH WITH DOT MOVED BELOW	NO	PVALID
ڙ	06A3	ARABIC LETTER FEH WITH DOT BELOW	NO	PVALID
ڙ	06A4	ARABIC LETTER VEH	NO	PVALID
ڙ	06A5	ARABIC LETTER FEH WITH THREE DOTS BELOW	NO	PVALID
ڙ	06A6	ARABIC LETTER PEHEH	NO	PVALID
ڙ	06A7	ARABIC LETTER QAF WITH	NO	PVALID

		DOT ABOVE		
ڦ	06A8	ARABIC LETTER QAF WITH THREE DOTS ABOVE	NO	PVALID
ڪ	06A9	ARABIC LETTER KEHEH	YES	PVALID
ڪا	06AA	ARABIC LETTER SWASH KAF	YES (variant of base character)	PVALID
ڪي	06AB	ARABIC LETTER KAF WITH RING	YES (variant of base character)	PVALID
ڪن	06AC	ARABIC LETTER KAF WITH DOT ABOVE	NO	PVALID
ڻ	06AD	ARABIC LETTER NG	NO	PVALID
ڳ	06AE	ARABIC LETTER KAF WITH THREE DOTS BELOW	NO	PVALID
گ	06AF	ARABIC LETTER GAF	YES	PVALID
گا	06B0	ARABIC LETTER GAF WITH RING	NO	PVALID
گي	06B1	ARABIC LETTER NGOEH	NO	PVALID
گي	06B2	ARABIC LETTER GAF WITH TWO DOTS BELOW	NO	PVALID
گي	06B3	ARABIC LETTER GUEH	NO	PVALID
گي	06B4	ARABIC LETTER GAF WITH THREE DOTS ABOVE	NO	PVALID
ڙ	06B5	ARABIC LETTER LAM WITH SMALL V	NO	PVALID
ڙ	06B6	ARABIC LETTER LAM WITH DOT ABOVE	NO	PVALID
ڙ	06B7	ARABIC LETTER LAM WITH THREE DOTS ABOVE	NO	PVALID
ڙ	06B8	ARABIC LETTER LAM WITH THREE DOTS BELOW	NO	PVALID
ڻ	06B9	ARABIC LETTER NOON WITH DOT BELOW	NO	PVALID
ڻ	06BA	ARABIC LETTER NOON GHUNNA	NO	PVALID
ڻ	06BB	ARABIC LETTER RNOON	NO	PVALID
ڻ	06BC	ARABIC LETTER NOON WITH RING	YES	PVALID
ڻ	06BD	ARABIC LETTER NOON WITH THREE DOTS ABOVE	NO	PVALID
ھ	06BE	ARABIC LETTER HEH DOACHASHMEE	YES (variant of base character)	PVALID
ھ	06BF	ARABIC LETTER TCHEH WITH DOT ABOVE	NO	PVALID
ھ	06C0	ARABIC LETTER HEH WITH YEHA ABOVE	YES	PVALID
ھ	06C1	ARABIC LETTER HEH GOAL	YES (variant of base character)	PVALID
ھ	06C2	ARABIC LETTER HEH GOAL WITH HAMZA ABOVE	YES (variant of base character)	PVALID
ھ	06C3	ARABIC LETTER TEH	YES (variant of	PVALID

		MARBUTA GOAL	base character)	
و	06C4	ARABIC LETTER WAW WITH RING	NO	PVALID
و	06C5	ARABIC LETTER KIRGHIZ OE	NO	PVALID
و	06C6	ARABIC LETTER OE	NO	PVALID
و	06C7	ARABIC LETTER U	NO	PVALID
و	06C8	ARABIC LETTER YU	NO	PVALID
و	06C9	ARABIC LETTER KIRGHIZ YU	NO	PVALID
و	06CA	ARABIC LETTER WAW WITH TWO DOTS ABOVE	NO	PVALID
و	06CB	ARABIC LETTER VE	NO	PVALID
ى	06CC	ARABIC LETTER FARSI YEH	YES (variant of base character)	PVALID
ى	06CD	ARABIC LETTER YEH WITH TAIL	YES	PVALID
ى	06CE	ARABIC LETTER YEH WITH SMALL V	NO	PVALID
و	06CF	ARABIC LETTER WAW WITH DOT ABOVE	NO	PVALID
ى	06D0	ARABIC LETTER E	YES	PVALID
ى	06D1	ARABIC LETTER YEH WITH THREE DOTS BELOW	NO	PVALID
ء	06D2	ARABIC LETTER YEH BARREE	YES (variant of base character)	PVALID
ء	06D3	ARABIC LETTER YEH BARREE WITH HAMZA ABOVE	YES (variant of base character)	PVALID
.	06D4	ARABIC FULL STOP	NO	DISALLOWED
ه	06D5	ARABIC LETTER AE	YES	PVALID
ﻮ	06D6	ARABIC SMALL HIGH LIGATURE SAD WITH LAM WITH	NO	PVALID
ﻮ	06D7	ARABIC SMALL HIGH LIGATURE QAF WITH LAM WITH	NO	PVALID
ﻮ	06D8	ARABIC SMALL HIGH MEEM INITIAL FORM	NO	PVALID
ﻮ	06D9	ARABIC SMALL HIGH LAM ALEF	NO	PVALID
ﻮ	06DA	ARABIC SMALL HIGH JEEM	NO	PVALID
ﻮ	06DB	ARABIC SMALL HIGH THREE DOTS	NO	PVALID
ﻮ	06DC	ARABIC SMALL HIGH SEEN	NO	PVALID
○	06DD	ARABIC END OF AYAH	NO	CONTEXTO
○	06DE	ARABIC START OF RUB EL HIZB	NO	DISALLOWED
○	06DF	ARABIC SMALL HIGH ROUNDED ZERO	NO	PVALID
○	06E0	ARABIC SMALL HIGH UPRIGHT RECTANGULAR ZERO	NO	PVALID

غ	06FC	ARABIC LETTER GHAIN WITH DOT BELOW	NO	PVALID
ء	06FD	ARABIC SIGN SINDHI AMPERSAND	NO	DISALLOWED
م	06FE	ARABIC SIGN SINDHI POSTPOSITION MEN	NO	DISALLOWED
ه	06FF	ARABIC LETTER HEH WITH INVERTED V	NO	PVALID
لله	FDF2	ARABIC LIGATURE ALLAH ISOLATED FORM	NO	DISALLOWED
اكبر	FDF3	ARABIC LIGATURE AKBAR ISOLATED FORM	NO	DISALLOWED
□	FDF4	ARABIC LIGATURE MOHAMMAD ISOLATED FORM	NO	DISALLOWED
بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ	FDF5	ARABIC LIGATURE SALAM ISOLATED FORM	NO	DISALLOWED
رسول	FDF6	ARABIC LIGATURE RASOUL ISOLATED FORM	NO	DISALLOWED
عليه	FDF7	ARABIC LIGATURE ALAYHE ISOLATED FORM	NO	DISALLOWED
بِالْحَمْدِ لِلَّهِ الرَّحْمَنِ الرَّحِيمِ	FDF8	ARABIC LIGATURE WASALLAM ISOLATED FORM	NO	DISALLOWED
ﷺ	FDF9	ARABIC LIGATURE SALLA ISOLATED FORM	NO	DISALLOWED
ﷺ ﷺ ﷺ	FDFA	ARABIC LIGATURE SALLALLAHOU ALAYHE WASALLAM	NO	DISALLOWED
□	FDFB	ARABIC LIGATURE JALLAJALALOUHOU	NO	DISALLOWED

Appendix A.8: Sinhala Character Set

Character	Unicode	Description	Language Community Decision	IDNAbis Decision[1]
	0D80	<reserved>		UNASSIGNED
	0D81	<reserved>		UNASSIGNED
◌◌	0D82	SINHALA SIGN ANUSVARAYA	YES	PVALID
◌◌◌	0D83	SINHALA SIGN VISARGAYA	YES	PVALID
	0D84	<reserved>		UNASSIGNED
ආ	0D85	SINHALA LETTER AYANNA	YES	PVALID
ආආ	0D86	SINHALA LETTER AAYANNA	YES	PVALID
ආඨ	0D87	SINHALA LETTER AEYANNA	YES	PVALID
ආඨඨ	0D88	SINHALA LETTER AEEYANNA	YES	PVALID
ඉ	0D89	SINHALA LETTER IYANNA	YES	PVALID

ඊ	0D8A	SINHALA LETTER IYANNA	YES	PVALID
උ	0D8B	SINHALA LETTER UYANNA	YES	PVALID
ඌ	0D8C	SINHALA LETTER UUYANNA	YES	PVALID
ඍ	0D8D	SINHALA LETTER IRUYANNA	YES	PVALID
ඎ	0D8E	SINHALA LETTER IRUUYANNA	YES	PVALID
ඏ	0D8F	SINHALA LETTER ILUYANNA	YES	PVALID
ඐ	0D90	SINHALA LETTER ILUUYANNA	YES	PVALID
එ	0D91	SINHALA LETTER EYANNA	YES	PVALID
ඒ	0D92	SINHALA LETTER EEEYANNA	YES	PVALID
ඓ	0D93	SINHALA LETTER AIYANNA	YES	PVALID
ඔ	0D94	SINHALA LETTER OYANNA	YES	PVALID
ඕ	0D95	SINHALA LETTER OOOYANNA	YES	PVALID
ඖ	0D96	SINHALA LETTER AUYANNA	YES	PVALID
	0D97	<reserved>		UNASSIGNED
	0D98	<reserved>		UNASSIGNED
	0D99	<reserved>		UNASSIGNED
ක	0D9A	SINHALA LETTER ALPAPRAANA KAYANNA	YES	PVALID
ඛ	0D9B	SINHALA LETTER MAHAAPRAANA KAYANNA	YES	PVALID
ග	0D9C	SINHALA LETTER ALPAPRAANA GAYANNA	YES	PVALID
ඝ	0D9D	SINHALA LETTER MAHAAPRAANA GAYANNA	YES	PVALID
ඞ	0D9E	SINHALA LETTER KANTAJA NAASIKYAYA	YES	PVALID
ඟ	0D9F	SINHALA LETTER SANYAKA GAYANNA	YES	PVALID
ච	0DA0	SINHALA LETTER ALPAPRAANA CAYANNA	YES	PVALID
ඡ	0DA1	SINHALA LETTER MAHAAPRAANA CAYANNA	YES	PVALID

ඊ	0DA2	SINHALA LETTER MAHAAPRAANA JAYANNA	YES	PVALID
ක	0DA3	SINHALA LETTER MAHAAPRAANA JAYANNA	YES	PVALID
කැ	0DA4	SINHALA LETTER TAALUJA NAASIKYAYA	YES	PVALID
කේ	0DA5	SINHALA LETTER TAALUJA SANYOOGA NAAKSIKYAYA	YES	PVALID
ඊ	0DA6	SINHALA LETTER SANYAKA JAYANNA	YES	PVALID
උ	0DA7	SINHALA LETTER ALPAPRAANA TTAYANNA	YES	PVALID
උ	0DA8	SINHALA LETTER MAHAAPRAANA TTAYANNA	YES	PVALID
උ	0DA9	SINHALA LETTER ALPAPRAANA DDAYANNA	YES	PVALID
උ	0DAA	SINHALA LETTER MAHAAPRAAN DDAYANNA	YES	PVALID
උ	0DAB	SINHALA LETTER MUURDHAJA NAYANNA	YES	PVALID
උ	0DAC	SINHALA LETTER SANYAKA DDAYANNA	YES	PVALID
උ	0DAD	SINHALA LETTER ALPAPRAANA TAYANNA	YES	PVALID
උ	0DAE	SINHALA LETTER MAHAAPRAANA TAYANNA	YES	PVALID
උ	0DAF	SINHALA LETTER ALPAPRAANA DAYANNA	YES	PVALID
උ	0DB0	SINHALA LETTER MAHAAPRAANA DAYANNA	YES	PVALID
උ	0DB1	SINHALA LETTER DANTAJA NAYANNA	YES	PVALID
	0DB2	<reserved>		UNASSIGNED
උ	0DB3	SINHALA LETTER SANYAKA DAYANNA	YES	PVALID
උ	0DB4	SINHALA LETTER ALPAPRAANA PAYANNA	YES	PVALID

ඵ	0DB5	SINHALA LETTER MAHAAPRAANA PAYANNA	YES	PVALID
ඵ	0DB6	SINHALA LETTER ALPAPRAANA BAYANNA	YES	PVALID
ඵ	0DB7	SINHALA LETTER MAHAAPRAANA BAYANNA	YES	PVALID
ඵ	0DB8	SINHALA LETTER MAYANNA	YES	PVALID
ඵ	0DB9	SINHALA LETTER AMBA BAYANNA	YES	PVALID
ඵ	0DBA	SINHALA LETTER YAYANNA	YES	PVALID
ඵ	0DBB	SINHALA LETTER RAYANNA	YES	PVALID
	0DBC	<reserved>		UNASSIGNED
ඵ	0DBD	SINHALA LETTER DANTAJA LAYANNA	YES	PVALID
	0DBE	<reserved>		UNASSIGNED
	0DBF	<reserved>		UNASSIGNED
ඵ	0DC0	SINHALA LETTER VAYANNA	YES	PVALID
ඵ	0DC1	SINHALA LETTER TAALUJA SAYANNA	YES	PVALID
ඵ	0DC2	SINHALA LETTER MUURDHAJA SAYANNA	YES	PVALID
ඵ	0DC3	SINHALA LETTER DANTAJA SAYANNA	YES	PVALID
ඵ	0DC4	SINHALA LETTER HAYANNA	YES	PVALID
ඵ	0DC5	SINHALA LETTER MUURDHAJA LAYANNA	YES	PVALID
ඵ	0DC6	SINHALA LETTER FAYANNA	YES	PVALID
	0DC7	<reserved>		UNASSIGNED
	0DC8	<reserved>		UNASSIGNED
	0DC9	<reserved>		UNASSIGNED
ඵ	0DCA	SINHALA SIGN AL-LAKUNA	YES	PVALID
	0DCB	<reserved>		UNASSIGNED
	0DCC	<reserved>		UNASSIGNED
	0DCD	<reserved>		UNASSIGNED
	0DCE	<reserved>		UNASSIGNED
ඵ	0DCF	SINHALA VOWEL SIGN AELA-PILLA	YES	PVALID

෧	0DD0	SINHALA VOWEL SIGN KETTI AEDAPILLA	YES	PVALID
෧	0DD1	SINHALA VOWEL SIGN DIGA AEDAPILLA	YES	PVALID
෧	0DD2	SINHALA VOWEL SIGN KETTI ISPILLA	YES	PVALID
෧	0DD3	SINHALA VOWEL SIGN DIGA IS- PILLA	YES	PVALID
෧	0DD4	SINHALA VOWEL SIGN KETTI PAAPILLA	YES	PVALID
	0DD5	<reserved>		UNASSIGNED
෧	0DD6	SINHALA VOWEL SIGN DIGA PAAPILLA	YES	PVALID
	0DD7	<reserved>		UNASSIGNED
෧	0DD8	SINHALA VOWEL SIGN GAETTAPILLA	YES	PVALID
෧	0DD9	SINHALA VOWEL SIGN KOMBUVA	YES	PVALID
෧	0DDA	SINHALA VOWEL SIGN DIGA KOMBUVA	YES	PVALID
෧	0DDB	SINHALA VOWEL SIGN KOMBU DEKA	YES	PVALID
෧	0DDC	SINHALA VOWEL SIGN KOMBUVA HAA AELA-PILLA	YES	PVALID
෧	0DDD	SINHALA VOWEL SIGN KOMBUVA HAA DIGA AELA-PILLA	YES	PVALID
෧	0DDE	SINHALA VOWEL SIGN KOMBUVA HA GAYANUKITTA	YES	PVALID
෧	0DDF	SINHALA VOWEL SIGN GAYANUKITTA	YES	PVALID
	0DE0	<reserved>		UNASSIGNED
	0DE1	<reserved>		UNASSIGNED
	0DE2	<reserved>		UNASSIGNED
	0DE3	<reserved>		UNASSIGNED
	0DE4	<reserved>		UNASSIGNED
	0DE5	<reserved>		UNASSIGNED
	0DE6	<reserved>		UNASSIGNED
	0DE7	<reserved>		UNASSIGNED
	0DE8	<reserved>		UNASSIGNED

	606	<reserved>		UNASSIGNED
	607	<reserved>		UNASSIGNED
	608	<reserved>		UNASSIGNED
	609	<reserved>		UNASSIGNED
	060A	<reserved>		UNASSIGNED
ف	060B	AFGHANI SIGN	NO	DISALLOWED
،	060C	ARABIC COMMA	NO	DISALLOWED
ر	060D	ARABIC DATE SEPARATOR	NO	DISALLOWED
ـ	060E	ARABIC POETIC VERSE SIGN	NO	DISALLOWED
ب	060F	ARABIC SIGN MISRA	NO	DISALLOWED
ر	610	ARABIC SIGN SALLALLAHOU ALAYHE WASSALLAM	YES	PVALID
ر	611	ARABIC SIGN ALAYHE ASSALLAM	YES	PVALID
ر	612	ARABIC SIGN RAHMATULLAH ALAYHE	YES	PVALID
ر	613	ARABIC SIGN RADI ALLAHOU ANHU	YES	PVALID
ر	614	ARABIC SIGN TAKHALLUS	YES	PVALID
□	615	ARABIC SMALL HIGH TAH	NO	PVALID
	616	<reserved>		UNASSIGNED
	617	<reserved>		UNASSIGNED
	618	<reserved>		UNASSIGNED
	619	<reserved>		UNASSIGNED
	061A	<reserved>		UNASSIGNED
؛	061B	ARABIC SEMICOLON	NO	DISALLOWED
		<reserved>		UNASSIGNED
		<reserved>		UNASSIGNED

⋮	061E	ARABIC TRIPLE DOT PUNCTUATION MARK	NO	DISALLOWED
؟	061F	ARABIC QUESTION MARK	NO	DISALLOWED
	620	<reserved>		UNASSIGNED
ء	621	ARABIC LETTER HAMZA	YES	PVALID
آ	622	ARABIC LETTER ALEF WITH MADDA ABOVE	YES	PVALID
أ	623	ARABIC LETTER ALEF WITH HAMZA ABOVE	YES	PVALID
ؤ	624	ARABIC LETTER WAW WITH HAMZA ABOVE	YES	PVALID
إ	625	ARABIC LETTER ALEF WITH HAMZA BELOW	YES (variant of base character)	PVALID
ئ	626	ARABIC LETTER YEH WITH HAMZA ABOVE	YES (variant of base character)	PVALID
ا	627	ARABIC LETTER ALEF	YES	PVALID
ب	628	ARABIC LETTER BEH	YES	PVALID
ة	629	ARABIC LETTER TEH MARBUTA	YES (variant of base character)	PVALID
ت	062A	ARABIC LETTER TEH	YES	PVALID
ث	062B	ARABIC LETTER THEH	YES	PVALID
ج	062C	ARABIC LETTER JEEM	YES	PVALID
ح	062D	ARABIC LETTER HAH	YES	PVALID
خ	062E	ARABIC LETTER KHAH	YES	PVALID
د	062F	ARABIC LETTER DAL	YES	PVALID
ذ	630	ARABIC LETTER THAL	YES	PVALID
ر	631	ARABIC LETTER REH	YES	PVALID
ز	632	ARABIC LETTER ZAIN	YES	PVALID
س	633	ARABIC LETTER SEEN	YES	PVALID
ش	634	ARABIC LETTER SHEEN	YES	PVALID
ص	635	ARABIC LETTER SAD	YES	PVALID
ض	636	ARABIC LETTER DAD	YES	PVALID
ط	637	ARABIC LETTER TAH	YES	PVALID
ظ	638	ARABIC LETTER ZAH	YES	PVALID
ع	639	ARABIC LETTER AIN	YES	PVALID
غ	063A	ARABIC LETTER GHAIN	YES	PVALID
	063B	<reserved>		UNASSIGNED
	063C	<reserved>		UNASSIGNED

	063D	<reserved>		UNASSIGNED
	063E	<reserved>		UNASSIGNED
	063F	<reserved>		UNASSIGNED
-	640	ARABIC TATWEEL	NO	PVALID
ف	641	ARABIC LETTER FEH	YES	PVALID
ق	642	ARABIC LETTER QAF	YES	PVALID
ك	643	ARABIC LETTER KAF	YES (variant of base character)	PVALID
ل	644	ARABIC LETTER LAM	YES	PVALID
م	645	ARABIC LETTER MEEM	YES	PVALID
ن	646	ARABIC LETTER NOON	YES	PVALID
ه	647	ARABIC LETTER HEH	YES (variant of base character)	PVALID
و	648	ARABIC LETTER WAW	YES	PVALID
ى	649	ARABIC LETTER ALEF MAKSURA	YES (variant of base character)	PVALID
ي	064A	ARABIC LETTER YEH	YES (variant of base character)	PVALID
َ	064B	ARABIC FATHATAN	YES	PVALID
ِ	064C	ARABIC DAMMATAN	YES	PVALID
ُ	064D	ARABIC KASRATAN	YES	PVALID
ا	064E	ARABIC FATHA	YES	PVALID
اَ	064F	ARABIC DAMMA	YES	PVALID
اِ	650	ARABIC KASRA	YES	PVALID

	651	ARABIC SHADDA	YES	PVALID
	652	ARABIC SUKUN	YES	PVALID
	653	ARABIC MADDAH ABOVE	YES	PVALID
	654	ARABIC HAMZA ABOVE	YES	PVALID
	655	ARABIC HAMZA BELOW	YES	PVALID
	656	ARABIC SUBSCRIPT ALEF	YES	PVALID
	657	ARABIC INVERTED DAMMA	NO	PVALID
	658	ARABIC MARK NOON GHUNNA	YES	PVALID
◌̣	659	ARABIC ZWARAKAY	NO	PVALID
◌̥	065A	ARABIC VOWEL SIGN SMALL V ABOVE	NO	PVALID
◌̦	065B	ARABIC VOWEL SIGN INVERTED SMALL V ABOVE	NO	PVALID
◌̧	065C	ARABIC VOWEL SIGN DOT BELOW	NO	PVALID
◌̨	065D	ARABIC REVERSED DAMMA	NO	PVALID
◌̩	065E	ARABIC FATHA WITH TWO DOTS	NO	PVALID
	065F	<reserved>		UNASSIGNED
۰	660	ARABIC-INDIC DIGIT ZERO	YES (variant of base character)	PVALID
۱	661	ARABIC-INDIC DIGIT ONE	YES (variant of base character)	PVALID
۲	662	ARABIC-INDIC DIGIT TWO	YES (variant of base character)	PVALID
۳	663	ARABIC-INDIC DIGIT THREE	YES (variant of base character)	PVALID

٤	664	ARABIC-INDIC DIGIT FOUR	YES (variant of base character)	PVALID
٥	665	ARABIC-INDIC DIGIT FIVE	YES (variant of base character)	PVALID
٦	666	ARABIC-INDIC DIGIT SIX	YES (variant of base character)	PVALID
٧	667	ARABIC-INDIC DIGIT SEVEN	YES (variant of base character)	PVALID
٨	668	ARABIC-INDIC DIGIT EIGHT	YES (variant of base character)	PVALID
٩	669	ARABIC-INDIC DIGIT NINE	YES (variant of base character)	PVALID
٪	066A	ARABIC PERCENT SIGN	NO	DISALLOWED
؄	066B	ARABIC DECIMAL SEPARATOR	NO	DISALLOWED
	066C	ARABIC THOUSANDS SEPARATOR	NO	DISALLOWED
*	066D	ARABIC FIVE POINTED STAR	NO	DISALLOWED
ﺀ	066E	ARABIC LETTER DOTLESS BEH	NO	PVALID
ﻖ	066F	ARABIC LETTER DOTLESS QAF	NO	PVALID
ﺀ	670	ARABIC LETTER SUPERSCRIPIT ALEF	YES	PVALID
ﺀ	671	ARABIC LETTER ALEF WASLA	NO	PVALID
ﺀ	672	ARABIC LETTER ALEF WITH WAVY HAMZA ABOVE	NO	PVALID
ﺀ	673	ARABIC LETTER ALEF WITH WAVY HAMZA BELOW	NO	PVALID
ﺀ	674	ARABIC LETTER HIGH HAMZA	NO	PVALID
ﺀ	675	ARABIC LETTER HIGH HAMZA ALEF	NO	DISALLOWED
ﺀ	676	ARABIC LETTER HIGH HAMZA WAW	NO	DISALLOWED
ﺀ	677	ARABIC LETTER U WITH HAMZA ABOVE	NO	DISALLOWED

ﻩ	678	ARABIC LETTER HIGH HAMZA YEH	NO	DISALLOWED
ﺕ	679	ARABIC LETTER TTEH	YES	PVALID
ﺕ	067A	ARABIC LETTER TTEHEH	NO	PVALID
ﺕ	067B	ARABIC LETTER BEEH	NO	PVALID
ﺕ	067C	ARABIC LETTER TEH WITH RING	NO	PVALID
ﺕ	067D	ARABIC LETTER TEH WITH THREE DOTS ABOVE DOWN	NO	PVALID
ﻩ	067E	ARABIC LETTER PEH	YES	PVALID
ﺕ	067F	ARABIC LETTER TEHEH	NO	PVALID
ﺕ	680	ARABIC LETTER BEHEH	NO	PVALID
ﺕ	681	ARABIC LETTER HAH WITH HAMZA ABOVE	NO	PVALID
ﺕ	682	ARABIC LETTER HAH WITH TWO DOTS VERTICAL ABO	NO	PVALID
ﻩ	683	ARABIC LETTER NYEH	NO	PVALID
ﻩ	684	ARABIC LETTER DYEH	NO	PVALID
ﺕ	685	ARABIC LETTER HAH WITH THREE DOTS ABOVE	NO	PVALID
ﺕ	686	ARABIC LETTER TCHEH	YES	PVALID
ﺕ	687	ARABIC LETTER TCHEHEH	NO	PVALID
ﺕ	688	ARABIC LETTER DDAL	YES	PVALID
ﺕ	689	ARABIC LETTER DAL WITH RING	NO	PVALID
ﺕ	068A	ARABIC LETTER DAL WITH DOT BELOW	NO	PVALID
ﺕ	068B	ARABIC LETTER DAL WITH DOT BELOW AND SMALL T	NO	PVALID
ﺕ	068C	ARABIC LETTER DAHAL	NO	PVALID
ﺕ	068D	ARABIC LETTER DDAHAL	NO	PVALID
ﺕ	068E	ARABIC LETTER DUL	NO	PVALID
ﺕ	068F	ARABIC LETTER DAL WITH THREE DOTS ABOVE DOWN	NO	PVALID
ﺕ	690	ARABIC LETTER DAL WITH FOUR DOTS ABOVE	NO	PVALID
ﺕ	691	ARABIC LETTER RREH	YES	PVALID
ﺕ	692	ARABIC LETTER REH WITH SMALL V	NO	PVALID
ﺕ	693	ARABIC LETTER REH WITH RING	NO	PVALID
ﺕ	694	ARABIC LETTER REH WITH DOT BELOW	NO	PVALID
ﺕ	695	ARABIC LETTER REH WITH SMALL V BELOW	NO	PVALID
ﺕ	696	ARABIC LETTER REH WITH DOT BELOW AND DOT ABO	NO	PVALID
ﺕ	697	ARABIC LETTER REH WITH TWO DOTS ABOVE	NO	PVALID

ژ	698	ARABIC LETTER JEH	YES	PVALID
ز	699	ARABIC LETTER REH WITH FOUR DOTS ABOVE	NO	PVALID
زى	069A	ARABIC LETTER SEEN WITH DOT BELOW AND DOT AB	NO	PVALID
زى	069B	ARABIC LETTER SEEN WITH THREE DOTS BELOW	NO	PVALID
زى	069C	ARABIC LETTER SEEN WITH THREE DOTS BELOW AND	NO	PVALID
زى	069D	ARABIC LETTER SAD WITH TWO DOTS BELOW	NO	PVALID
زى	069E	ARABIC LETTER SAD WITH THREE DOTS ABOVE	NO	PVALID
ظ	069F	ARABIC LETTER TAH WITH THREE DOTS ABOVE	NO	PVALID
ع	06A0	ARABIC LETTER AIN WITH THREE DOTS ABOVE	NO	PVALID
و	06A1	ARABIC LETTER DOTLESS FEH	NO	PVALID
و	06A2	ARABIC LETTER FEH WITH DOT MOVED BELOW	NO	PVALID
ف	06A3	ARABIC LETTER FEH WITH DOT BELOW	NO	PVALID
ف	06A4	ARABIC LETTER VEH	NO	PVALID
ف	06A5	ARABIC LETTER FEH WITH THREE DOTS BELOW	NO	PVALID
ق	06A6	ARABIC LETTER PEHEH	NO	PVALID
ق	06A7	ARABIC LETTER QAF WITH DOT ABOVE	NO	PVALID
ق	06A8	ARABIC LETTER QAF WITH THREE DOTS ABOVE	NO	PVALID
ك	06A9	ARABIC LETTER KEHEH	YES	PVALID
ك	06AA	ARABIC LETTER SWASH KAF	NO	PVALID
ك	06AB	ARABIC LETTER KAF WITH RING	NO	PVALID
ك	06AC	ARABIC LETTER KAF WITH DOT ABOVE	NO	PVALID
ك	06AD	ARABIC LETTER NG	NO	PVALID
ك	06AE	ARABIC LETTER KAF WITH THREE DOTS BELOW	NO	PVALID
گ	06AF	ARABIC LETTER GAF	YES	PVALID
گ	06B0	ARABIC LETTER GAF WITH RING	NO	PVALID
گ	06B1	ARABIC LETTER NGOEH	NO	PVALID
گ	06B2	ARABIC LETTER GAF WITH TWO DOTS BELOW	NO	PVALID
گ	06B3	ARABIC LETTER GUEH	NO	PVALID
گ	06B4	ARABIC LETTER GAF WITH THREE DOTS ABOVE	NO	PVALID
ل	06B5	ARABIC LETTER LAM WITH SMALL V	NO	PVALID

ل	06B6	ARABIC LETTER LAM WITH DOT ABOVE	NO	PVALID
ل	06B7	ARABIC LETTER LAM WITH THREE DOTS ABOVE	NO	PVALID
ل	06B8	ARABIC LETTER LAM WITH THREE DOTS BELOW	NO	PVALID
ن	06B9	ARABIC LETTER NOON WITH DOT BELOW	NO	PVALID
ن	06BA	ARABIC LETTER NOON GHUNNA	YES	PVALID
ن	06BB	ARABIC LETTER RNOON	NO	PVALID
ن	06BC	ARABIC LETTER NOON WITH RING	NO	PVALID
ن	06BD	ARABIC LETTER NOON WITH THREE DOTS ABOVE	NO	PVALID
ه	06BE	ARABIC LETTER HEH DOACHASHMEE	YES	PVALID
ه	06BF	ARABIC LETTER TCHEH WITH DOT ABOVE	NO	PVALID
ه	06C0	ARABIC LETTER HEH WITH YEH ABOVE	YES (variant of base character)	PVALID
ه	06C1	ARABIC LETTER HEH GOAL	YES	PVALID
ه	06C2	ARABIC LETTER HEH GOAL WITH HAMZA ABOVE	YES	PVALID
ه	06C3	ARABIC LETTER TEH MARBUTA GOAL	YES	PVALID
و	06C4	ARABIC LETTER WAW WITH RING	NO	PVALID
و	06C5	ARABIC LETTER KIRGHIZ OE	NO	PVALID
و	06C6	ARABIC LETTER OE	NO	PVALID
و	06C7	ARABIC LETTER U	NO	PVALID
و	06C8	ARABIC LETTER YU	NO	PVALID
و	06C9	ARABIC LETTER KIRGHIZ YU	NO	PVALID
و	06CA	ARABIC LETTER WAW WITH TWO DOTS ABOVE	NO	PVALID
و	06CB	ARABIC LETTER VE	NO	PVALID
ى	06CC	ARABIC LETTER FARSI YEH	YES	PVALID
ى	06CD	ARABIC LETTER YEH WITH TAIL	NO	PVALID
ى	06CE	ARABIC LETTER YEH WITH SMALL V	NO	PVALID
و	06CF	ARABIC LETTER WAW WITH DOT ABOVE	NO	PVALID
ى	06D0	ARABIC LETTER E	NO	PVALID
ى	06D1	ARABIC LETTER YEH WITH THREE DOTS BELOW	NO	PVALID
ى	06D2	ARABIC LETTER YEH BARREE	YES	PVALID

ء	06D3	ARABIC LETTER YEH BARREE WITH HAMZA ABOVE	YES	PVALID
-	06D4	ARABIC FULL STOP	NO	DISALLOWED
ه	06D5	ARABIC LETTER AE	NO	PVALID
ص	06D6	ARABIC SMALL HIGH LIGATURE SAD WITH LAM WITH	NO	PVALID
ق	06D7	ARABIC SMALL HIGH LIGATURE QAF WITH LAM WITH	NO	PVALID
م	06D8	ARABIC SMALL HIGH MEEM INITIAL FORM	NO	PVALID
لا	06D9	ARABIC SMALL HIGH LAM ALEF	NO	PVALID
ج	06DA	ARABIC SMALL HIGH JEEM	NO	PVALID
ث	06DB	ARABIC SMALL HIGH THREE DOTS	NO	PVALID
س	06DC	ARABIC SMALL HIGH SEEN	NO	PVALID
○	06DD	ARABIC END OF AYAH	NO	CONTEXTO
⦿	06DE	ARABIC START OF RUB EL HIZB	NO	DISALLOWED
◌◌◌	06DF	ARABIC SMALL HIGH ROUNDED ZERO	NO	PVALID
◌◌◌	06E0	ARABIC SMALL HIGH UPRIGHT RECTANGULAR ZERO	NO	PVALID
◌◌◌	06E1	ARABIC SMALL HIGH DOTLESS HEAD OF KHAH	NO	PVALID
م	06E2	ARABIC SMALL HIGH MEEM ISOLATED FORM	NO	PVALID
س	06E3	ARABIC SMALL LOW SEEN	NO	PVALID
◌◌◌	06E4	ARABIC SMALL HIGH MADDA	NO	PVALID
و	06E5	ARABIC SMALL WAW	NO	PVALID
ي	06E6	ARABIC SMALL YEH	NO	PVALID
◌◌◌	06E7	ARABIC SMALL HIGH YEH	NO	PVALID
ن	06E8	ARABIC SMALL HIGH NOON	NO	PVALID
↑	06E9	ARABIC PLACE OF SAJDAH	NO	DISALLOWED
◌◌◌	06EA	ARABIC EMPTY CENTRE LOW STOP	NO	PVALID
◌◌◌	06EB	ARABIC EMPTY CENTRE HIGH STOP	NO	PVALID
◌◌◌	06EC	ARABIC ROUNDED HIGH STOP WITH FILLED CENTRE	NO	PVALID
م	06ED	ARABIC SMALL LOW MEEM	NO	PVALID
ذ	06EE	ARABIC LETTER DAL WITH INVERTED V	NO	PVALID
ر	06EF	ARABIC LETTER REH WITH	NO	PVALID

		INVERTED V		
٠	06F0	EXTENDED ARABIC-INDIC DIGIT ZERO	YES	PVALID
١	06F1	EXTENDED ARABIC-INDIC DIGIT ONE	YES	PVALID
٢	06F2	EXTENDED ARABIC-INDIC DIGIT TWO	YES	PVALID
٣	06F3	EXTENDED ARABIC-INDIC DIGIT THREE	YES	PVALID
٤	06F4	EXTENDED ARABIC-INDIC DIGIT FOUR	YES	PVALID
٥	06F5	EXTENDED ARABIC-INDIC DIGIT FIVE	YES	PVALID
٦	06F6	EXTENDED ARABIC-INDIC DIGIT SIX	YES	PVALID
٧	06F7	EXTENDED ARABIC-INDIC DIGIT SEVEN	YES	PVALID
٨	06F8	EXTENDED ARABIC-INDIC DIGIT EIGHT	YES	PVALID
٩	06F9	EXTENDED ARABIC-INDIC DIGIT NINE	YES	PVALID
ش	06FA	ARABIC LETTER SHEEN WITH DOT BELOW	NO	PVALID
ض	06FB	ARABIC LETTER DAD WITH DOT BELOW	NO	PVALID
غ	06FC	ARABIC LETTER GHAIN WITH DOT BELOW	NO	PVALID
ء	06FD	ARABIC SIGN SINDHI AMPERSAND	NO	DISALLOWED
م	06FE	ARABIC SIGN SINDHI POSTPOSITION MEN	NO	DISALLOWED
هـ	06FF	ARABIC LETTER HEH WITH INVERTED V	NO	PVALID
الله	FDF2	ARABIC LIGATURE ALLAH ISOLATED FORM	NO	DISALLOWED
اكبر	FDF3	ARABIC LIGATURE AKBAR ISOLATED FORM	NO	DISALLOWED
□	FDF4	ARABIC LIGATURE MOHAMMAD ISOLATED FORM	NO	DISALLOWED
بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ	FDF5	ARABIC LIGATURE SALAM ISOLATED FORM	NO	DISALLOWED
رسول	FDF6	ARABIC LIGATURE RASOUL ISOLATED FORM	NO	DISALLOWED
عليه	FDF7	ARABIC LIGATURE ALAYHE ISOLATED FORM	NO	DISALLOWED

بِسْمِ الرَّحْمَنِ الرَّحِيمِ	FDF8	ARABIC LIGATURE WASALLAM ISOLATED FORM	NO	DISALLOWED
ﷺ	FDF9	ARABIC LIGATURE SALLA ISOLATED FORM	NO	DISALLOWED
اللَّهُمَّ صَلِّ وَسَلِّمْ وَاغْنِنِي مِنَ الْيَقْوَمِ	FDFA	ARABIC LIGATURE SALLALLAHOU ALAYHE WASALLAM	NO	DISALLOWED
□	FDFB	ARABIC LIGATURE JALLAJALALOUHOU	NO	DISALLOWED