

Title:	Comments concerning Unicode Security considerations (draft TR36)
Source:	Michel Suignard, Microsoft
Action:	Review by UTC

### **List 1: List of allowable IDN characters on input**

The following is a list of Unicode characters representing what should be allowable IDN characters on input. They may be transformed by the IDN transformation (Nameprep) in another character sequences from that same list. This is especially true for upper case characters and some special case such as ß: 00DF LATIN SMALL LETTER SHARP S which is transformed into 'ss'.

The list has been validated against the lists published by the DE NIC at <http://www.denic.de/en/domains/idns/liste.html>, the JP NIC at <http://www.iana.org/assignments/idn/jp-japanese.html>, the KR NIC at <http://www.iana.org/assignments/idn/kr-korean.html>, the Chinese NICs at <http://www.ietf.org/internet-drafts/draft-xdlee-idn-cdnadmin-03.txt> and other sources at <http://www.iana.org/assignments/idn/registered.htm>. As more information is found, it will be updated. The list is by definition a subset of allowed input characters to the IDN transformation functions as defined by the IDN RFCs.

Good source of info are <http://about.museum/idn/language.html> and <http://www.iana.org/assignments/idn/>.

The list is by no mean a final proposal. The concept is to use as an input to further reduce the allowable list based on confusable characters. For example, it is very likely that allowable characters from the UNIFIED CANADIAN ABORIGINAL SYLLABICS block will be significantly reduced.

The general principle was to extend the DNS LDH (Latin, Digit, Hyphen) principle to IDN, basically restricting the IDN repertoire to letters, no symbol, with as little exception as possible.

002D	; HYPHEN-MINUS	048A-04CE	; CYRILLIC
0030-0039	; DIGIT	04D0-04F5	; CYRILLIC
0041-005A	; UPPERCASE LATIN ALPHABET	04F8-04F9	; CYRILLIC
0061-007A	; LOWERCASE LATIN ALPHABET	0500-050F	; CYRILLIC SUPPLEMENTARY
00C0-00D6	; LATIN-1 LETTERS	0531-0556	; ARMENIAN
00D8-00F6	; LATIN-1 LETTERS	0561-0586	; ARMENIAN
00F8-00FF	; LATIN-1 LETTERS	0591-05A1	; HEBREW COMBINING
0100-0131	; LATIN EXTENDED	05A3-05B9	; HEBREW COMBINING
0134-013E	; LATIN EXTENDED	05BB-05BD	; HEBREW COMBINING
0141-0148	; LATIN EXTENDED	05BF	; HEBREW COMBINING
014A-017E	; LATIN EXTENDED	05C1-05C2	; HEBREW COMBINING
0180-01C3	; LATIN EXTENDED	05C4	; HEBREW COMBINING
01CD-01F0	; LATIN EXTENDED	05D0-05EA	; HEBREW LETTER
01F4-0220	; LATIN EXTENDED	05F0-05F2	; HEBREW LIGATURE
0222-0233	; LATIN EXTENDED	0621-063A	; ARABIC LETTER
0250-02AD	; IPA EXTENSIONS	0641-0655	; ARABIC LETTER
0300-033F	; COMBINING DIACRITICAL MARKS	0660-0669	; ARABIC DIGIT
0342	; COMBINING DIACRITICAL MARKS	066E-0674	; ARABIC LETTER
0345-034F	; COMBINING DIACRITICAL MARKS	0679-06D3	; ARABIC LETTER
0360-036F	; COMBINING DIACRITICAL MARKS	06D5	; ARABIC LETTER
0386	; GREEK	06D6-06DC	; ARABIC ANNOTATION
0388-038A	; GREEK	06DF-06E8	; ARABIC ANNOTATION
038C	; GREEK	06EA-06ED	; ARABIC ANNOTATION
038E-03A1	; GREEK	06F0-06FC	; ARABIC EXTENDED
03A3-03CE	; GREEK	0710-072C	; SYRIAC
03D7-03EF	; GREEK	0730-074A	; SYRIAC
03F3	; GREEK	0780-07B1	; THAANA
0400-0481	; CYRILLIC	0901-0903	; DEVANAGARI SIGN
0483-0486	; CYRILLIC	0905-0939	; DEVANAGARI LETTER

093C-094D	; DEVANAGARI SIGN	0BD7	; TAMIL SIGN
0950-0954	; DEVANAGARI SIGN	0BE7-0BEF	; TAMIL DIGIT
0960-0963	; DEVANAGARI ADDITION	0C01-0C03	; TELUGU SIGN
0966-096F	; DEVANAGARI DIGIT	0C05-0C0C	; TELUGU VOWEL
0981-0983	; BENGALI SIGN	0C0E-0C10	; TELUGU VOWEL
0985-098C	; BENGALI VOWEL	0C12-0C28	; TELUGU VOWEL AND CONSONANT
098F-0990	; BENGALI VOWEL	0C2A-0C33	; TELUGU CONSONANT
0993-09A8	; BENGALI CONSONANT	0C35-0C39	; TELUGU CONSONANT
09AA-09B0	; BENGALI CONSONANT	0C3E-0C44	; TELUGU SIGN
09B2	; BENGALI CONSONANT	0C46-0C48	; TELUGU SIGN
09B6-09B9	; BENGALI CONSONANT	0C4A-0C4D	; TELUGU SIGN
09BC	; BENGALI SIGN	0C55-0C56	; TELUGU SIGN
09BE-09C4	; BENGALI SIGN	0C60-0C61	; TELUGU ADDITIONS
09C7-09C8	; BENGALI SIGN	0C66-0C6F	; TELUGU DIGIT
09CB-09CD	; BENGALI SIGN	0C82-0C83	; KANNADA SIGN
09D7	; BENGALI SIGN	0C85-0C8C	; KANNADA VOWEL
09E0-09E3	; BENGALI ADDITION	0C8E-0C90	; KANNADA VOWEL
09E6-09F1	; BENGALI DIGIT AND LETTER	0C92-0CA8	; KANNADA VOWEL AND CONSONANT
0A02	; GURMUKHI SIGN	0CAA-0CB3	; KANNADA CONSONANT
0A05-0A0A	; GURMUKHI VOWEL	0CB5-0CB9	; KANNADA CONSONANT
0A0F-0A10	; GURMUKHI VOWEL	0CBE-0CC4	; KANNADA SIGN
0A13-0A28	; GURMUKHI VOWEL AND CONSONANT	0CC6-0CC8	; KANNADA SIGN
0A2A-0A30	; GURMUKHI CONSONANT	0CCA-0CCD	; KANNADA SIGN
0A32-0A32	; GURMUKHI CONSONANT	0CD5-0CD6	; KANNADA SIGN
0A35-0A35	; GURMUKHI CONSONANT	0CDE	; KANNADA CONSONANT
0A38-0A39	; GURMUKHI CONSONANT	0CE0-0CE1	; KANNADA ADDITION
0A3C	; GURMUKHI SIGN	0CE6-0CEF	; KANNADA DIGIT
0A3E-0A42	; GURMUKHI SIGN	0D02-0D03	; MALAYALAM SIGN
0A47-0A48	; GURMUKHI SIGN	0D05-0D0C	; MALAYALAM VOWEL
0A4B-0A4D	; GURMUKHI SIGN	0D0E-0D10	; MALAYALAM VOWEL
0A5C	; GURMUKHI CONSONANT	0D12-0D28	; MALAYALAM VOWEL AND CONSONANT
0A66-0A74	; GURMUKHI DIGIT AND ADDITION	0D2A-0D39	; MALAYALAM CONSONANT
0A81-0A83	; GUJARATI SIGN	0D3E-0D43	; MALAYALAM SIGN
0A85-0A8B	; GUJARATI VOWEL	0D46-0D48	; MALAYALAM SIGN
0A8D	; GUJARATI VOWEL	0D4A-0D4D	; MALAYALAM SIGN
0A8F-0A91	; GUJARATI VOWEL	0D57	; MALAYALAM SIGN
0A93-0AA8	; GUJARATI VOWEL AND CONSONANT	0D60-0D61	; MALAYALAM ADDITION
0AAA-0AB0	; GUJARATI CONSONANT	0D66-0D6F	; MALAYALAM DIGIT
0AB2-0AB3	; GUJARATI CONSONANT	0D82-0D83	; SINHALA SIGN
0AB5-0AB9	; GUJARATI CONSONANT	0D85-0D96	; SINHALA VOWEL
0ABC-0AC5	; GUJARATI SIGN	0D9A-0DB1	; SINHALA CONSONANT
0AC7-0AC9	; GUJARATI SIGN	0DB3-0DBB	; SINHALA CONSONANT
0ACB-0ACD	; GUJARATI SIGN	0DBD	; SINHALA CONSONANT
0AE0	; GUJARATI ADDITION	0DC0-0DC6	; SINHALA CONSONANT
0AE6-0AEF	; GUJARATI DIGIT	0DCA	; SINHALA SIGN
0B01-0B03	; ORIYA SIGN	0DCF-0DD4	; SINHALA SIGN
0B05-0B0C	; ORIYA VOWEL	0DD6	; SINHALA SIGN
0B0F-0B10	; ORIYA VOWEL	0DD8-0DDF	; SINHALA SIGN
0B13-0B28	; ORIYA VOWEL AND CONSONANT	0DF2-0DF3	; SINHALA SIGN
0B2A-0B30	; ORIYA CONSONANT	0E01-0E32	; THAI CONSONANT, SIGN AND VOWEL
0B32-0B33	; ORIYA CONSONANT	0E34-0E3A	; THAI VOWEL
0B36-0B39	; ORIYA CONSONANT	0E40-0E4E	; THAI VOWEL, SIGN AND TONE MARK
0B3C-0B43	; ORIYA SIGN	0E50-0E59	; THAI DIGIT
0B47-0B48	; ORIYA SIGN	0E81-0E82	; LAO CONSONANT
0B4B-0B4D	; ORIYA SIGN	0E84	; LAO CONSONANT
0B56-0B57	; ORIYA SIGN	0E87-0E88	; LAO CONSONANT
0B5F-0B61	; ORIYA CONSONANT AND ADDITION	0E8A	; LAO CONSONANT
0B66-0B6F	; ORIYA DIGIT	0E8D	; LAO CONSONANT
0B82-0B83	; TAMIL SIGN	0E94-0E97	; LAO CONSONANT
0B85-0B8A	; TAMIL VOWEL	0E99-0E9F	; LAO CONSONANT
0B8E-0B90	; TAMIL VOWEL	0EA1-0EA3	; LAO CONSONANT
0B92-0B95	; TAMIL VOWEL AND CONSONANT	0EA5	; LAO CONSONANT
0B99-0B9A	; TAMIL CONSONANT	0EA7	; LAO CONSONANT
0B9C	; TAMIL CONSONANT	0EAA-0EAB	; LAO CONSONANT
0B9E-0B9F	; TAMIL CONSONANT	0EAD-0EB2	; LAO CONSONANT, SIGN AND VOWEL
0BA3-0BA4	; TAMIL CONSONANT	0EB4-0EB9	; LAO VOWEL
0BA8-0BAA	; TAMIL CONSONANT	0EBB-0EBD	; LAO VOWEL, SIGN
0BAE-0BB5	; TAMIL CONSONANT	0EC0-0EC4	; LAO VOWEL
0BB7-0BB9	; TAMIL CONSONANT	0EC6	; LAO SIGN
0BBE-0BC2	; TAMIL SIGN	0EC8-0ECD	; LAO TONE MARK AND SIGN
0BC6-0BC8	; TAMIL SIGN	0ED0-0ED9	; LAO DIGIT
0BCA-0BCD	; TAMIL SIGN	0F00	; TIBETAN SYLLABLE

0F18-0F19	; TIBETAN SIGN	1760-176C	; TAGBANWA
0F20-0F29	; TIBETAN DIGIT	176E-1770	; TAGBANWA
0F35	; TIBETAN SIGN	1772-1773	; TAGBANWA
0F37	; TIBETAN SIGN	1780-17B3	; KHMER
0F39	; TIBETAN SIGN	17B6-17D2	; KHMER
0F3E-0F42	; TIBETAN SIGN AND CONSONANT	17D7	; KHMER
0F44-0F47	; TIBETAN CONSONANT	17DC	; KHMER
0F49-0F4C	; TIBETAN CONSONANT	17E0-17E9	; KHMER
0F4E-0F51	; TIBETAN CONSONANT	1810-1819	; MONGOLIAN
0F53-0F56	; TIBETAN CONSONANT	1820-1877	; MONGOLIAN
0F58-0F5B	; TIBETAN CONSONANT	1880-18A9	; MONGOLIAN
0F5D-0F68	; TIBETAN CONSONANT	1E00-1E99	; LATIN EXTENDED ADDITIONAL
0F6A	; TIBETAN CONSONANT	1EA0-1EF9	; LATIN EXTENDED ADDITIONAL
0F71-0F72	; TIBETAN VOWEL	1F00-1F15	; GREEK EXTENDED
0F74	; TIBETAN VOWEL	1F18-1F1D	; GREEK EXTENDED
0F7A-0F80	; TIBETAN	1F20-1F45	; GREEK EXTENDED
0F82-0F8B	; TIBETAN	1F48-1F4D	; GREEK EXTENDED
0F90-0F92	; TIBETAN SUBJOINED CONSONANT	1F50-1F57	; GREEK EXTENDED
0F94-0F97	; TIBETAN SUBJOINED CONSONANT	1F59	; GREEK EXTENDED
0F99-0F9C	; TIBETAN SUBJOINED CONSONANT	1F5B	; GREEK EXTENDED
0F9E-0FA1	; TIBETAN SUBJOINED CONSONANT	1F5D	; GREEK EXTENDED
0FA3-0FA6	; TIBETAN SUBJOINED CONSONANT	1F5F-1F70	; GREEK EXTENDED
0FA8-0FAB	; TIBETAN SUBJOINED CONSONANT	1F72	; GREEK EXTENDED
0FAD-0FB8	; TIBETAN SUBJOINED CONSONANT	1F74	; GREEK EXTENDED
0FBA-0FBC	; TIBETAN SUBJOINED CONSONANT	1F76	; GREEK EXTENDED
1000-1021	; MYANMAR CONSONANT AND VOWEL	1F78	; GREEK EXTENDED
1023-1027	; MYANMAR VOWEL	1F7A	; GREEK EXTENDED
1029-102A	; MYANMAR VOWEL	1F7C	; GREEK EXTENDED
102C-1032	; MYANMAR VOWEL	1F80-1FB4	; GREEK EXTENDED
1036-1039	; MYANMAR SIGN	1FB6-1FBA	; GREEK EXTENDED
1040-1049	; MYANMAR DIGIT	1FBC	; GREEK EXTENDED
1050-1059	; MYANMAR EXTENSION	1FC2-1FC4	; GREEK EXTENDED
10A0-10C5	; GEORGIAN KHUTSURI	1FC6-1FC8	; GREEK EXTENDED
10D0-10F8	; GEORGIAN MKHEDRULI AND OTHER	1FCA	; GREEK EXTENDED
1100-1159	; HANGUL JAMO	1FCC	; GREEK EXTENDED
115F-11A2	; HANGUL JAMO	1FD0-1FD2	; GREEK EXTENDED
11A8-11F9	; HANGUL JAMO	1FD6-1FDA	; GREEK EXTENDED
1200-1206	; ETHIOPIIC SYLLABLE	1FE0-1FE2	; GREEK EXTENDED
1208-1246	; ETHIOPIIC SYLLABLE	1FE4-1FEA	; GREEK EXTENDED
1248	; ETHIOPIIC SYLLABLE	1FEC	; GREEK EXTENDED
124A-124D	; ETHIOPIIC SYLLABLE	1FF2-1FF4	; GREEK EXTENDED
1250-1256	; ETHIOPIIC SYLLABLE	1FF6-1FF8	; GREEK EXTENDED
1258	; ETHIOPIIC SYLLABLE	1FFA	; GREEK EXTENDED
125A-125D	; ETHIOPIIC SYLLABLE	1FFC	; GREEK EXTENDED
1260-1286	; ETHIOPIIC SYLLABLE	2019	; RIGHT SINGLE QUOTATION MARK
1288	; ETHIOPIIC SYLLABLE	2800-28FF	; BRAILLE PATTERNS
128A-128D	; ETHIOPIIC SYLLABLE	3003	; DITTO MARK
1290-12AE	; ETHIOPIIC SYLLABLE	3005-3007	; IDEOGRAPHIC MARKS
12B0	; ETHIOPIIC SYLLABLE	3041-3096	; HIRAGANA
12B2-12B5	; ETHIOPIIC SYLLABLE	3099-309A	; HIRAGANA
12B8-12BE	; ETHIOPIIC SYLLABLE	309D-309E	; HIRAGANA
12C0	; ETHIOPIIC SYLLABLE	30A1-30FE	; KATAKANA
12C2-12C5	; ETHIOPIIC SYLLABLE	3105-312C	; BOPOMOFO
12C8-12CE	; ETHIOPIIC SYLLABLE	31A0-31B7	; BOPOMOFO EXTENDED
12D0-12D6	; ETHIOPIIC SYLLABLE	31F0-31FF	; KATAKANA PHONETIC EXTENSIONS
12D8-12EE	; ETHIOPIIC SYLLABLE	3400-4DB5	; CJK UNIFIED IDEOGRAPHS EXT A
12F0-130E	; ETHIOPIIC SYLLABLE	4E00-9FA5	; CJK UNIFIED IDEOGRAPHS
1310	; ETHIOPIIC SYLLABLE	A000-A48C	; YI SYLLABLES
1312-1315	; ETHIOPIIC SYLLABLE	AC00-D7A3	; HANGUL SYLLABLES
1318-131E	; ETHIOPIIC SYLLABLE	FA0E-FA0F	; CJK COMPATIBILITY IDEOGRAPHS
1320-1346	; ETHIOPIIC SYLLABLE	FA11	; CJK COMPATIBILITY IDEOGRAPHS
1348-135A	; ETHIOPIIC SYLLABLE	FA13-FA14	; CJK COMPATIBILITY IDEOGRAPHS
1369-1371	; ETHIOPIIC DIGIT	FA1F	; CJK COMPATIBILITY IDEOGRAPHS
13A0-13F4	; CHEROKEE	FA21	; CJK COMPATIBILITY IDEOGRAPHS
1401-166C	; UNIFIED CANADIAN ABORIGINAL SYL	FA23-FA24	; CJK COMPATIBILITY IDEOGRAPHS
166F-1676	; UNIFIED CANADIAN ABORIGINAL SYL	FA27-FA29	; CJK COMPATIBILITY IDEOGRAPHS
1681-169A	; OGHAM	10300-1031E	; OLD ITALIC
16A0-16EA	; RUNIC	10330-1034A	; GOTHIC
1700-170C	; TAGALOG	10400-10425	; DESERET
170E-1714	; TAGALOG	10428-1044D	; DESERET
1720-1734	; HANUNOO	20000-2A6D6	; CJK UNIFIED IDEOGRAPHS EXT B
1740-1753	; BUHID		

**List 2: List of allowable IDN characters that cannot be first character of an IDN label**

This list was constructed by extracting characters from the previous list that were `XID_Continue_only` and were not digits (`GC <> Nd`).

0300-033F	; COMBINING DIACRITICAL MARKS	0C3E-0C44	; TELUGU SIGN
0342	; COMBINING DIACRITICAL MARKS	0C46-0C48	; TELUGU SIGN
0345-034F	; COMBINING DIACRITICAL MARKS	0C4A-0C4D	; TELUGU SIGN
0360-036F	; COMBINING DIACRITICAL MARKS	0C55-0C56	; TELUGU SIGN
0483-0486	; CYRILLIC COMBINING	0C82-0C83	; KANNADA SIGN
0591-05A1	; HEBREW COMBINING	0CBE-0CC4	; KANNADA SIGN
05A3-05B9	; HEBREW COMBINING	0CC6-0CC8	; KANNADA SIGN
05BB-05BD	; HEBREW COMBINING	0CCA-0CCD	; KANNADA SIGN
05BF	; HEBREW COMBINING	0CD5-0CD6	; KANNADA SIGN
05C1-05C2	; HEBREW COMBINING	0D02-0D03	; MALAYALAM SIGN
05C4	; HEBREW COMBINING	0D3E-0D43	; MALAYALAM SIGN
064B-0655	; ARABIC COMBINING	0D46-0D48	; MALAYALAM SIGN
0670	; ARABIC COMBINING	0D4A-0D4D	; MALAYALAM SIGN
06D6-06DC	; ARABIC ANNOTATION	0D57	; MALAYALAM SIGN
06DF-06E4	; ARABIC COMBINING	0D82-0D83	; SINHALA SIGN
06E7-06E8	; ARABIC COMBINING	0DCA	; SINHALA SIGN
06EA-06ED	; ARABIC ANNOTATION	0DCF-0DD4	; SINHALA SIGN
0711	; SYRIAC COMBINING	0DD6	; SINHALA SIGN
0730-074A	; SYRIAC COMBINING	0DD8-0DDF	; SINHALA SIGN
07A6-07B0	; THAANA COMBINING	0DF2-0DF3	; SINHALA SIGN
0901-0903	; DEVAAGARI SIGN	0E31	; THAI VOWEL
093C	; DEVANAGARI SIGN	0E34-0E3A	; THAI VOWEL
093E-094D	; DEVANAGARI SIGN	0E47-0E4E	; THAI VOWEL, SIGN AND TONE MARK
0951-0954	; DEVANAGARI SIGN	0EB1	; LAO VOWEL
0962-0963	; DEVANAGARI ADDITION	0EB4-0EB9	; LAO VOWEL
0981-0983	; BENGALI SIGN	0EBB-0EBC	; LAO VOWEL, SIGN
09BC	; BENGALI SIGN	0EC8-0ECD	; LAO TONE MARK AND SIGN
09BE-09C4	; BENGALI SIGN	0F18-0F19	; TIBETAN SIGN
09C7-09C8	; BENGALI SIGN	0F35	; TIBETAN SIGN
09CB-09CD	; BENGALI SIGN	0F37	; TIBETAN SIGN
09D7	; BENGALI SIGN	0F39	; TIBETAN SIGN
09E2-09E3	; BENGALI ADDITION	0F3E-0F3F	; TIBETAN SIGN
0A02	; GURMUKHI SIGN	0F71-0F72	; TIBETAN VOWEL
0A3C	; GURMUKHI SIGN	0F74	; TIBETAN VOWEL
0A3E-0A42	; GURMUKHI SIGN	0F7A-0F80	; TIBETAN
0A47-0A48	; GURMUKHI SIGN	0F82-0F84	; TIBETAN MARK AND SIGN
0A4B-0A4D	; GURMUKHI SIGN	0F86-0F87	; TIBETAN MARK AND SIGN
0A70-0A71	; GURMUKHI ADDITION	0F90-0F92	; TIBETAN SUBJOINED CONSONANT
0A81-0A83	; GUJARATI SIGN	0F94-0F97	; TIBETAN SUBJOINED CONSONANT
0ABC	; GUJARATI SIGN	0F99-0F9C	; TIBETAN SUBJOINED CONSONANT
0ABE-0AC5	; GUJARATI SIGN	0F9E-0FA1	; TIBETAN SUBJOINED CONSONANT
0AC7-0AC9	; GUJARATI SIGN	0FA3-0FA6	; TIBETAN SUBJOINED CONSONANT
0ACB-0ACD	; GUJARATI SIGN	0FA8-0FAB	; TIBETAN SUBJOINED CONSONANT
0B01-0B03	; ORIYA SIGN	0FAD-0FB8	; TIBETAN SUBJOINED CONSONANT
0B3C	; ORIYA SIGN	0FBA-0FBC	; TIBETAN SUBJOINED CONSONANT
0B3E-0B43	; ORIYA SIGN	102C-1032	; MYANMAR VOWEL
0B47-0B48	; ORIYA SIGN	1036-1039	; MYANMAR SIGN
0B4B-0B4D	; ORIYA SIGN	1056-1059	; MYANMAR EXTENSION
0B56-0B57	; ORIYA SIGN	1712-1714	; TAGALOG SIGN
0B82	; TAMIL SIGN	1732-1734	; HANUNOO SIGN
0BBE-0BC2	; TAMIL SIGN	1752-1753	; BUHID SIGN
0BC6-0BC8	; TAMIL SIGN	1772-1773	; TAGBANWA SIGN
0BCA-0BCD	; TAMIL SIGN	17B6-17D2	; KHMER
0BD7	; TAMIL SIGN	18A9	; MONGOLIAN
0C01-0C03	; TELUGU SIGN	3099-309A	; HIRAGANA

### **List 3: List of allowable IDN characters that are not XID Continue**

These characters were added for various reasons. 002D HYPHEN-MINUS is there because it is an essential part of DNS host name. 0F85 TIBETAN MARK PALUTA is there because it is related to the various Avagraha signs which are part of XID Continue. 2019 RIGHT SINGLE QUOTATION MARK is there to express the apostrophe. It isn't clear why 2800-28FF BRAILLE PATTERNS should be excluded from IDN. Finally, 3003 DITTO MARK and 30FB KATAKANA MIDDLE DOT have been added because they are part of the IDN collection created by the JP tld registry.

002D	; HYPHEN-MINUS	2800-28FF	; BRAILLE PATTERNS
0F85	; TIBETAN MARK PALUTA	3003	; DITTO MARK
2019	; RIGHT SINGLE QUOTATION MARK	30FB	; KATAKANA MIDDLE DOT

### **List 4: Characters that are Identifiers (XID Continue) & NFKD but not part of the IDN subset**

02B9	Lm;MODIFIER LETTER PRIME
02BA	Lm;MODIFIER LETTER DOUBLE PRIME
02BB	Lm;MODIFIER LETTER TURNED COMMA
02BC	Lm;MODIFIER LETTER APOSTROPHE
02BD	Lm;MODIFIER LETTER REVERSED COMMA
02BE	Lm;MODIFIER LETTER RIGHT HALF RING
02BF	Lm;MODIFIER LETTER LEFT HALF RING
02C0	Lm;MODIFIER LETTER GLOTTAL STOP
02C1	Lm;MODIFIER LETTER REVERSED GLOTTAL STOP
02C6	Lm;MODIFIER LETTER CIRCUMFLEX ACCENT
02C7	Lm;CARON
02C8	Lm;MODIFIER LETTER VERTICAL LINE
02C9	Lm;MODIFIER LETTER MACRON
02CA	Lm;MODIFIER LETTER ACUTE ACCENT
02CB	Lm;MODIFIER LETTER GRAVE ACCENT
02CC	Lm;MODIFIER LETTER LOW VERTICAL LINE
02CD	Lm;MODIFIER LETTER LOW MACRON
02CE	Lm;MODIFIER LETTER LOW GRAVE ACCENT
02CF	Lm;MODIFIER LETTER LOW ACUTE ACCENT
02D0	Lm;MODIFIER LETTER TRIANGULAR COLON
02D1	Lm;MODIFIER LETTER HALF TRIANGULAR COLON
02EE	Lm;MODIFIER LETTER DOUBLE APOSTROPHE
0559	Lm;ARMENIAN MODIFIER LETTER LEFT HALF RING
0640	Lm;ARABIC TATWEEL
0AD0	Lo;GUJARATI OM
16EE	Nl;RUNIC ARLAUG SYMBOL
16EF	Nl;RUNIC TVIMADUR SYMBOL
16F0	Nl;RUNIC BELGTHOR SYMBOL
2118	So;SCRIPT CAPITAL P
212E	So;ESTIMATED SYMBOL
2180	Nl;ROMAN NUMERAL ONE THOUSAND C D
2181	Nl;ROMAN NUMERAL FIVE THOUSAND
2182	Nl;ROMAN NUMERAL TEN THOUSAND
2183	Nl;ROMAN NUMERAL REVERSED ONE HUNDRED
3021	Nl;HANGZHOU NUMERAL ONE
3022	Nl;HANGZHOU NUMERAL TWO
3023	Nl;HANGZHOU NUMERAL THREE
3024	Nl;HANGZHOU NUMERAL FOUR
3025	Nl;HANGZHOU NUMERAL FIVE
3026	Nl;HANGZHOU NUMERAL SIX
3027	Nl;HANGZHOU NUMERAL SEVEN
3028	Nl;HANGZHOU NUMERAL EIGHT
3029	Nl;HANGZHOU NUMERAL NINE
3031	Lm;VERTICAL KANA REPEAT MARK
3032	Lm;VERTICAL KANA REPEAT WITH VOICED SOUND MARK
3033	Lm;VERTICAL KANA REPEAT MARK UPPER HALF
3034	Lm;VERTICAL KANA REPEAT WITH VOICED SOUND MARK UPPER HALF
3035	Lm;VERTICAL KANA REPEAT MARK LOWER HALF
303B	Lm;VERTICAL IDEOGRAPHIC ITERATION MARK
303C	Lo;MASU MARK
FE73	Lo;ARABIC TAIL FRAGMENT

**List 5: Characters that are XID Continue Only & NFKD but not part of the IDN subset**

005F Pc;LOW LINE  
00B7 Po;MIDDLE DOT  
0FC6 Mn;TIBETAN SYMBOL PADMA GDAN  
17D3 Mn;KHMER SIGN BATHAMASAT  
180B Mn;MONGOLIAN FREE VARIATION SELECTOR ONE  
180C Mn;MONGOLIAN FREE VARIATION SELECTOR TWO  
180D Mn;MONGOLIAN FREE VARIATION SELECTOR THREE  
203F Pc;UNDERTIE  
2040 Pc;CHARACTER TIE  
20D0 Mn;COMBINING LEFT HARPOON ABOVE  
20D1 Mn;COMBINING RIGHT HARPOON ABOVE  
20D2 Mn;COMBINING LONG VERTICAL LINE OVERLAY  
20D3 Mn;COMBINING SHORT VERTICAL LINE OVERLAY  
20D4 Mn;COMBINING ANTICLOCKWISE ARROW ABOVE  
20D5 Mn;COMBINING CLOCKWISE ARROW ABOVE  
20D6 Mn;COMBINING LEFT ARROW ABOVE  
20D7 Mn;COMBINING RIGHT ARROW ABOVE  
20D8 Mn;COMBINING RING OVERLAY  
20D9 Mn;COMBINING CLOCKWISE RING OVERLAY  
20DA Mn;COMBINING ANTICLOCKWISE RING OVERLAY  
20DB Mn;COMBINING THREE DOTS ABOVE  
20DC Mn;COMBINING FOUR DOTS ABOVE  
20E1 Mn;COMBINING LEFT RIGHT ARROW ABOVE  
20E5 Mn;COMBINING REVERSE SOLIDUS OVERLAY  
20E6 Mn;COMBINING DOUBLE VERTICAL STROKE OVERLAY  
20E7 Mn;COMBINING ANNUITY SYMBOL  
20E8 Mn;COMBINING TRIPLE UNDERDOT  
20E9 Mn;COMBINING WIDE BRIDGE ABOVE  
20EA Mn;COMBINING LEFTWARDS ARROW OVERLAY  
302A Mn;IDEOGRAPHIC LEVEL TONE MARK  
302B Mn;IDEOGRAPHIC RISING TONE MARK  
302C Mn;IDEOGRAPHIC DEPARTING TONE MARK  
302D Mn;IDEOGRAPHIC ENTERING TONE MARK  
302E Mn;HANGUL SINGLE DOT TONE MARK  
302F Mn;HANGUL DOUBLE DOT TONE MARK  
FB1E Mn;HEBREW POINT JUDEO-SPANISH VARIKA  
FE00 Mn;VARIATION SELECTOR-1  
FE01 Mn;VARIATION SELECTOR-2  
FE02 Mn;VARIATION SELECTOR-3  
FE03 Mn;VARIATION SELECTOR-4  
FE04 Mn;VARIATION SELECTOR-5  
FE05 Mn;VARIATION SELECTOR-6  
FE06 Mn;VARIATION SELECTOR-7  
FE07 Mn;VARIATION SELECTOR-8  
FE08 Mn;VARIATION SELECTOR-9  
FE09 Mn;VARIATION SELECTOR-10  
FE0A Mn;VARIATION SELECTOR-11  
FE0B Mn;VARIATION SELECTOR-12  
FE0C Mn;VARIATION SELECTOR-13  
FE0D Mn;VARIATION SELECTOR-14  
FE0E Mn;VARIATION SELECTOR-15  
FE0F Mn;VARIATION SELECTOR-16  
FE20 Mn;COMBINING LIGATURE LEFT HALF  
FE21 Mn;COMBINING LIGATURE RIGHT HALF  
FE22 Mn;COMBINING DOUBLE TILDE LEFT HALF  
FE23 Mn;COMBINING DOUBLE TILDE RIGHT HALF  
1D165 Mc;MUSICAL SYMBOL COMBINING STEM  
1D166 Mc;MUSICAL SYMBOL COMBINING SPRECHGESANG STEM  
1D167 Mn;MUSICAL SYMBOL COMBINING TREMOLO-1  
1D168 Mn;MUSICAL SYMBOL COMBINING TREMOLO-2  
1D169 Mn;MUSICAL SYMBOL COMBINING TREMOLO-3  
1D16D Mc;MUSICAL SYMBOL COMBINING AUGMENTATION DOT  
1D16E Mc;MUSICAL SYMBOL COMBINING FLAG-1  
1D16F Mc;MUSICAL SYMBOL COMBINING FLAG-2  
1D170 Mc;MUSICAL SYMBOL COMBINING FLAG-3  
1D171 Mc;MUSICAL SYMBOL COMBINING FLAG-4  
1D172 Mc;MUSICAL SYMBOL COMBINING FLAG-5  
1D17B Mn;MUSICAL SYMBOL COMBINING ACCENT

1D17C	Mn;	MUSICAL	SYMBOL	COMBINING	STACCATO
1D17D	Mn;	MUSICAL	SYMBOL	COMBINING	TENUTO
1D17E	Mn;	MUSICAL	SYMBOL	COMBINING	STACCATISSIMO
1D17F	Mn;	MUSICAL	SYMBOL	COMBINING	MARCATO
1D180	Mn;	MUSICAL	SYMBOL	COMBINING	MARCATO-STACCATO
1D181	Mn;	MUSICAL	SYMBOL	COMBINING	ACCENT-STACCATO
1D182	Mn;	MUSICAL	SYMBOL	COMBINING	LOURE
1D185	Mn;	MUSICAL	SYMBOL	COMBINING	DOIT
1D186	Mn;	MUSICAL	SYMBOL	COMBINING	RIP
1D187	Mn;	MUSICAL	SYMBOL	COMBINING	FLIP
1D188	Mn;	MUSICAL	SYMBOL	COMBINING	SMEAR
1D189	Mn;	MUSICAL	SYMBOL	COMBINING	BEND
1D18A	Mn;	MUSICAL	SYMBOL	COMBINING	DOUBLE TONGUE
1D18B	Mn;	MUSICAL	SYMBOL	COMBINING	TRIPLE TONGUE
1D1AA	Mn;	MUSICAL	SYMBOL	COMBINING	DOWN BOW
1D1AB	Mn;	MUSICAL	SYMBOL	COMBINING	UP BOW
1D1AC	Mn;	MUSICAL	SYMBOL	COMBINING	HARMONIC
1D1AD	Mn;	MUSICAL	SYMBOL	COMBINING	SNAP PIZZICATO

### **List 6: List of confusable characters**

At this point the confusable character information (see TR36 for further details on the confusable principle) is presented using a set of mapping tables with one character mapping into another character or a sequence of characters.

It is expected that the end result will be a mix of IDN character list reduction and a reduced set of mapping. The mapping tables only addresses on input characters that are part of the IDN list (there are still few additional symbols which should eventually be removed, depending on resolution about symbol inclusion in the IDN character list).

Mapping is only provided for characters that are preprocessed by NFKD, to eliminate all pre-composed characters.

### **Mapping definition**

The mapping tables have been separated in the following categories:

1. In within script confusable, lower case (when applicable)
2. In within script confusable (for non bicameral script)
3. In mixed script (script x – script y) confusable, lower case (when applicable)
4. In mixed script (script x – script y) confusable (for non bicameral script)
5. In within script confusable, upper case (when applicable)
6. In mixed script (script x – script y) confusable, upper case (when applicable)

These categories have been created to facilitate the analysis; eventually they could be rearranged to be closer to an implementation.

All tables have an input and an output character (or sequence). At this stage the output is not guaranteed to be final, i.e. some cascading may take place. Cascading situations have not been fully explored. Because of these cascading effects, the final characters may be quite different from the original, even more if an upper to lower case folding occurs during the cascading.

For example, any output of a confusable which is an upper case character has to be case-mapped according to the IDN transformation rules and then processed again through the confusable tables used by lower case characters.

The final result is a string of characters that may or may not be rendered in an acceptable manner by the layout system. Its primary purpose is to be a hash code for comparison purposes.

Note: Reducing the number of characters allowed by the List 1 would simplify the mapping stage, especially if some problematic digraphs are removed.

Establishing these tables raised the following issues:

- Although some work was done to match symbols with confusable letters, that part is far from complete. At this point, these mapping can be mostly ignored because most of these symbols are not valid IDN input characters (according to List 1).
- Output characters should never be a character not allowed in IDN according to list 1. Any such mapping should be ignored.
- There are multiple levels of confusability. In some case the glyphs are exactly identical (such as the Latin A and the Greek Alpha: Α). In many others, the glyphs are different but confusable at low resolution.
- Some scripts are inherently unsafe, based on their list of in-script confusable characters. A special mention must be given to the Canadian Aboriginal Syllabics which has well over 200 in-script confusable entries, mostly created by decomposable digraphs. The Latin script with over 40 entries in the same category is a distant second. Restricting further the List 1 to reduce these cases would make a lot of sense.
- Behind the in-script confusable threat, the AnyScript-Latin mix is also a common threat. This can be reduced significantly by blocking such mixes.
- These tables are still preliminary, more characters will be added, and the format will be changed to accommodate implementation needs.
- There is no intent to seek confusable among the 70 000 or so CJK unified ideographs. It is expected that concerned registries will have a conservative policy on that domain.
- Although typically real fonts are used to show the similarities, there are few cases when the representative glyphs are fake ones. This is mostly due to the use of preliminary fonts.
- The Canadian Aboriginal Syllabics deserve a normalization step to remove the duplicate encodings.

E.g. the sequence ᑭᐅ (1550 CANADIAN SYLLABICS R, 146C CANADIAN SYLLABICS KAAI) would be normalized into ᑭᐅ (157E CANADIAN SYLLABICS QAAI).

From looking at the table, the normalization steps can easily be constructed. It is necessary to normalize to the shortest form to avoid a complex layout engine to render these scripts. This is in the context of a further IDN input character restriction. In the context of establishing a confusable mapping function, the longest form is preferable.

The input value of the tables is postmarked with a ‘\*’ when an equivalent entry was found in TR36. When conflicting mapping exist, the various choices are shown. In many cases, the choice is between preciseness which brings little benefit in spoofing mitigation and less precise matching which captures many more threats.



**In Latin script confusable, lower case**

Input	Output	Display	Comment
0031 *	006C	1 ; l	Digit one -> Latin small letter l
00A5 *	0059 0335	Ÿ; Ȳ	Yen sign -> Latin capital letter y + Combining short stroke overlay
00E6	0061 0065	æ ; æ	Latin small letter ae -> Latin small letter a + Latin small letter e
00F8 *	006F 0337	ø ; ø	Latin small letter o with stroke -> Latin small letter o + Combining
0111 *	0064 0335	đ ; đ	Latin small letter d with stroke -> Latin small letter d + Combining short stroke overlay
0127 *	0068 0335	ħ ; ħ	Latin small letter h with stroke -> Latin small letter h + Combining short stroke overlay
0131 *	0069	ı ; i	Latin small letter dotless i -> Latin small letter i
0142 *	006C 0337	ł ; ł	Latin small letter with bar -> Latin small letter l + Combining short solidus overlay
0153	006F 0065	œ ; œ	Latin small ligature oe -> Latin small letter o + Latin small letter e
0167 *	0074 0335	ţ ; ţ	Latin small letter t with stroke -> Latin small letter t + Combining short stroke overlay
0180 *	0062 0335	ḃ ; ḃ	Latin small letter b with stroke -> Latin small letter b + Combining short stroke overlay
0183 *	0062 0304	Ḅ ; Ḅ	Latin small letter b with topbar -> Latin small letter b + Combining macron (also confusable with 0182 Latin capital letter b with topbar)
0184	0062	ḅ ; b	Latin capital letter tone six -> Latin small letter b
0185 *	0062	Ḇ ; b	Latin small letter tone six -> Latin small letter b
018C *	0064 0304	ḍ ; ḍ	Latin small letter d with topbar -> Latin small letter d + Combining macron
0192	0066	ƒ ; f	Latin small letter f with hook -> Latin small letter f
0192 *	0066 0322	ƒ ; ƒ	Latin small letter f with hook -> Latin small letter f + Combining retroflex below
0196 *	006C	ι ; l	Latin capital letter iota -> Latin small letter l
0199	006B	ƙ ; k	Latin small letter k with hook -> Latin small letter k
0199 *	006B 0302	ƙ ; ƙ	Latin small letter k with hook -> Latin small letter k + Combining circumflex accent
019A	0142	ł ; ł	Latin small letter l with bar -> Latin small letter l with stroke
019A *	006C 0335	ł ; ł	Latin small letter l with bar -> Latin small letter l + Combining short stroke overlay
019E	014B	ŋ ; η	Latin small letter n with long right leg -> Latin small letter eng
019E *	006E 0329	ŋ ; η	Latin small letter n with long right leg -> Latin small letter n + Combining vertical line below
01A5 *	0070 02C4	ṑ ; p^	Latin small letter p with hook -> latin small letter p + modifier letter up arrowhead

01AB	0074 0327	ṭ ; ṭ̸	Latin small letter t with palatal hook -> Latin small letter t + Combining cedilla
01AB *	0074 0321	ṭ̸ ; ṭ̸̸	Latin small letter t with palatal hook -> Latin small letter t + Combining palatalized hook below
01AD *	0074 0302	t̥ ; t̸̥	Latin small letter t with hook -> Latin small letter t + Combining circumflex accent
01B4	0079	ỵ ; ỵ	Latin small letter y with hook -> Latin small letter y
01B4 *	0079 0302	ỵ̸ ; ỵ̸̸	Latin small letter y with hook -> Latin small letter y + Combining circumflex accent
01B6	007A	ẓ ; ẓ	Latin small letter with stroke -> Latin small letter z
01B6 *	007A 0335	ẓ̸ ; ẓ̸̸	Latin small letter with stroke -> Latin small letter z + Combining short stroke overlay
01BD	0035	₅ ; ₅	Latin small letter tone five -> Digit five
01C0 *	007C	ḷ ; ḷ	Latin letter dental click -> Vertical line
01DD *	0259	ə̣ ; ə̣	Latin small letter turned e -> Latin small letter schwa
01E5	0067	g̣ ; g̣	Latin small letter g with stroke -> Latin small letter g
01E5 *	0067 0335	g̸̣ ; g̸̸̣	Latin small letter g with stroke -> Latin small letter g + Combining short stroke overlay
0223	0038	8̣ ; 8̣	Latin small letter ou -> Digit eight
0225 *	007A 0321	Ẓ ; Ẓ̸	Latin small letter z with hook -> Latin small letter z + Combining palatalized hook below
0251 *	0061	ɑ̣ ; ɑ̣	Latin small letter alpha -> Latin small letter a
0252 *	006F	ọ ; ọ	Latin small letter turned alpha ; Latin small letter o
0253	0062	ḅ ; ḅ	Latin small letter b with hook -> Latin small letter b
0253 *	0062 0302	ḅ̸ ; ḅ̸̸	Latin small letter b with hook -> Latin small letter b + Combining circumflex accent
0256	0064	ḍ ; ḍ	Latin small letter d with tail -> Latin small letter d
0256 *	0064 0322	ḍ̸ ; ḍ̸̸	Latin small letter d with tail -> Latin small letter d + Combining retroflex hook below
0257	0064	ɖ̣ ; ɖ̣	Latin small letter d with hook -> Latin small letter d
025A *	01DD 02DE	ə̸̣ ; ə̸̸̣	Latin small letter schwa with hook -> Latin small letter turned e + Modifier letter rhotic hook
0260	0067	g̸̣ ; g̸̣	Latin small letter g with hook -> Latin small letter g
0261 *	0067	g̸̸̣ ; g̸̸̸̣	Latin small letter script g -> Latin small letter g
0266	0068	ḥ ; ḥ	Latin small letter h with hook -> Latin small letter h

0266 *	0068 0302	h̃ ; h̃	Latin small letter h with hook -> Latin small letter h + Combining circumflex accent
0268 *	0069 0335	i̇ ; i̇	Latin small letter i with stroke -> Latin small letter i + Combining short stroke overlay
0269 *	0131	ı̇ ; ı̇	Latin small letter iota -> Latin small letter dotless i (TR36 says 0069)
026A *	0131	İ ; İ	Latin letter small capital I -> Latin small letter dotless i (TR36 says 0069)
026B	0142	l̃ ; l̃	Latin small letter l with middle tilde -> Latin small letter l with stroke
026B *	006C 0334	l̃ ; l̃	Latin small letter l with middle tilde -> Latin small letter l + Combining tilde overlay
026C	0142	l̃ ; l̃	Latin small letter l with belt -> Latin small letter l with stroke
026D	006C	ḷ ; ḷ	Latin small letter l with hook -> Latin small letter l
026D *	006C 0322	ḷ ; ḷ	Latin small letter l with hook -> Latin small letter l + Combining retroflex hook below
0271	006d	ṃ ; ṃ	Latin small letter m with hook -> Latin small letter m
0271 *	006d 0321	ṃ ; ṃ	Latin small letter m with hook -> Latin small letter m + Combining palatalized hook below
0272 *	006E 0321	ṇ ; ṇ	Latin small letter n with left hook -> Latin small letter n + Combining palatalized hook below
0273	014B	ɳ ; ɳ	Latin small letter n with retroflex hook -> Latin small letter eng
0273 *	006E 0322	ɳ ; ɳ	Latin small letter n with retroflex hook -> Latin small letter n + Combining retroflex hook below
027C *	0072 0329	ṛ ; ṛ	Latin small letter r with long leg -> Latin small letter r + Combining vertical line below
027D *	0072 0322	ṛ ; ṛ	Latin small letter r with tail -> Latin small letter r + Combining retroflex hook below
0282 *	0073 0322	ṣ ; ṣ	Latin small letter s with hook -> Latin small letter s + Combining retroflex hook below
0290 *	007A 0322	ẓ ; ẓ	Latin small letter z with retroflex hook -> Latin small letter z + Combining retroflex hook below
0292 *	01BA 0321	ʒ̣ ; ʒ̣	Latin small letter ezh -> Latin small letter ezh with tail + Combining palatalized hook below (????)
029D	006A	j̣ ; j̣	Latin small letter j with cross-tail -> Latin small letter j
02A0 *	0071 02C4	q̣ ; q̣	Latin small letter q with hook -> latin small letter q + modifier letter up arrowhead
02A3	0064 007A	dz ; dz	Latin small letter dz digraph -> Latin small letter d + Latin small letter z
02A4	0064 0292	dʒ ; dʒ	Latin small letter dezhdigraph -> Latin small letter d + Latin small letter ezh
02A5	0064 0291	dʒ̣ ; dʒ̣	Latin small letter dz digraph with curl -> Latin small d + Latin small letter z with curl
02A6	0074 0073	ts ; ts	Latin small letter ts digraph -> Latin small letter t + Latin small letter s

02A7	0074 0283	ʈʃ ; ʈʃ	Latin small letter tesh digraph -> Latin small letter t + Latin small letter esh
02A8	0074 0255	ʈʂ ; ʈʂ	Latin small letter tc digraph with curl -> Latin small letter t + Latin small letter c with curl
02A9	0066 014B	fŋ ; fŋ	Latin small letter feng -> Latin small letter f + Latin small letter eng
02AA	006C 0073	lʂ ; lʂ	Latin small letter ls digraph -> Latin small letter l + Latin small letter s
02AB	006C 007A	lʐ ; lʐ	Latin small letter lz digraph -> Latin small letter l + Latin small letter z
0326	0327	ˌ ; ˌ	Combining comma below – Combining cedilla
0326 *	0321	ˌ ; ˌ	Combining comma below – Combining palatalized hook below
0311 *	0302	˘ ; ˘	Combining inverted breve -> Combining circumflex accent
030C	0306	ˇ ; ˇ	Combining caron -> Combining breve

#### In Greek script confusable, lower case

Input	Output	Display	Comment
03C2	03DB	ς ; ς	Greek small letter final sigma -> Greek small letter stigma
2296	03B8	⊖ ; θ	Circled minus -> Greek small letter theta
236C	03B8	⊖ ; θ	APL functional symbol zilde -> Greek small letter theta
2373 *	03B9	ι ; ι	APL functional symbol iota -> Greek small letter iota (also confusable with 0131 Latin small letter dotless i)
2374 *	03C1	ρ ; ρ	APL functional symbol rho -> Greek small letter rho (also confusable with 0070 Latin small letter p)
2375	03C9	ω ; ω	APL functional symbol omega -> Greek small letter omega
237A *	03B1	α ; α	APL functional symbol alpha -> Greek small letter alpha (also confusable with 0061 Latin small letter a)

#### In Cyrillic script confusable, lower case

Input	Output	Display	Comment
0431 *	0036	б ; 6	Cyrillic small letter be -> Digit six
044D	0437	э ; э	Cyrillic small letter e -> Cyrillic small letter ze
045B *	0068 0335	ћ ; Һ	Cyrillic small letter tshe -> Latin small letter h + Combining short stroke overlay (also confusable with 0127 Latin small letter h with stroke)
0473	04E9	о̅ ; о̅	Cyrillic small letter fita -> Cyrillic small letter barred o (also confusable with 0275 Latin small letter barred o)
047D *	0461 0483	Ѡ ; Ѡ	Cyrillic small letter omega with titlo -> Cyrillic small letter omega + Combining Cyrillic titlo
048D *	0463	ѣ ; ѣ	Cyrillic small letter semisoft sign -> Cyrillic small letter yat

0493 *	0072 0335	ƚ ; ƚ	Cyrillic small letter r -> Latin small letter r + Combining short stroke overlay
0497 *	0436 0329	Ӏ ; Ӏ	Cyrillic small letter zhe with descender -> Cyrillic small letter zhe + Combining vertical line below
0499 *	025C 0321	ӑ ; ӑ	Cyrillic small letter ze with descender -> Latin small letter reversed open e + Combining palatalized hook below
049B	0138 0329	Ӌ ; Ӌ	Cyrillic small letter ka with descender -> Latin small letter kra + Combining vertical line below
049F	0138 0335	ӆ ; ӆ	Cyrillic small letter ka with stroke -> Latin small letter kra + Combining short stroke overlay
04A3 *	029C 0329	Ӈ ; Ӈ	Cyrillic small letter en -> Latin letter small capital h + Combining vertical line below
04AB	0063 0327	ӈ ; ӈ	Cyrillic small letter es with descender -> Latin small letter c + Combining cedilla
04AB *	0063 0322	ӈ ; ӈ	Cyrillic small letter es with descender -> Latin small letter c + Combining retroflex hook below
04AD *	03C4 0329	ӏ ; ӏ	Cyrillic small letter te -> Greek small letter tau + Combining vertical line below
04B1	00A5	ӑ ; ӑ	Cyrillic small letter straight u with stroke -> Yen sign
04B1 *	04AF 0335	ӑ ; ӑ	Cyrillic small letter straight u with stroke -> Cyrillic small letter straight u + Combining short stroke overlay
04B3 *	0078 0329	ӓ ; ӓ	Cyrillic small letter ha with descender -> Latin small letter x + Combining vertical line below
04BF *	04BD 0322	ӕ ; ӕ	Cyrillic small letter abkhasian che with descender -> Cyrillic small letter abkhasian che + Combining retroflex hook below
04C0 *	0456	Ӏ ; Ӏ	Cyrillic letter palochka -> Cyrillic small letter I (special case because 04C0 is case-invariant) (TR36 says 0049, same thing)
04C6 *	043B 0321	ӑ ; ӑ	Cyrillic small letter el with tail -> Cyrillic small letter el + Combining palatalized hook below
04C8 *	029C 0321	Ӈ ; Ӈ	Cyrillic small letter en with hook -> Latin letter small capital h + Combining palatalized hook below
04CA *	029C 0321	Ӈ ; Ӈ	Cyrillic small letter en with tail -> Latin letter small capital h + Combining palatalized hook below
04CC *	04B7	ӕ ; ӕ	Cyrillic small letter khakassian che -> Cyrillic small letter che with descender
04CE *	043C 0321	ӑ ; ӑ	Cyrillic small letter em with tail -> Cyrillic small letter em + Combining palatalized hook below

#### In Armenian script confusable, lower case

Input	Output	Display	Comment
0563	0566	զ ; զ	Armenian small letter gim -> Armenian small letter za (also confusable with 0071 Latin small q)
057C	0578	ր ; ռ	Armenian small letter ra -> Armenian small letter vo

#### In Arabic script confusable

Input	Output	Display	Comment
06D5 *	0647	ه ; ه	Arabic letter ae -> Arabic letter heh

06CC *	0649	ﻯ ; ﻯ	Arabic letter farsi yeh -> Arabic letter alef maksura
0674 *	0654	ﺀ ; ﺀ	Arabic letter high Hamza -> Arabic hamza above (combining)
06EC *	06DF	ﻻ ; ﻻ	Arabic rounded high stop with filled center ->Arabic small high rounded zero
06F0 *	0660	ﻮ ; ﻮ	Extended unavu-indic digit zero ->Arabic-indic digit zero
06F1 *	0661	ﻱ ; ﻱ	Extended unavu-indic digit one ->Arabic-indic digit one
06F2 *	0662	ﻮ ; ﻮ	Extended unavu-indic digit two ->Arabic-indic digit two
06F3 *	0663	ﻮ ; ﻮ	Extended unavu-indic digit three ->Arabic-indic digit three
06F5	0665	ﻮ ; ﻮ	Extended unavu-indic digit five ->Arabic-indic digit five
06F6	0621	ﻮ ; ﻮ	Extended unavu-indic digit six ->Arabic letter hamza (different directionality)
06F7 *	0667	ﻮ ; ﻮ	Extended unavu-indic digit seven ->Arabic-indic digit seven
06F8 *	0668	ﻮ ; ﻮ	Extended unavu-indic digit eight ->Arabic-indic digit eight
06F9 *	0669	ﻮ ; ﻮ	Extended unavu-indic digit nine ->Arabic-indic digit nine

#### In Devanagari script confusable

Input	Output	Display	Comment
0906 *	0905 093E	आ ; आ	Devanagari letter aa -> Devanagari letter a + Devanagari vowel sign aa
0911	0905 0949	आँ ; आँ	Devanagari letter candra o -> Devanagari letter a + Devanagari vowel sign candra o
0912	0905 094A	आो ; आो	Devanagari letter short o -> Devanagari letter a + Devanagari vowel sign short o
0913	0905 094B	आो ; आो	Devanagari letter o -> Devanagari letter a + Devanagari vowel sign o
0914	0905 094C	आौ ; आौ	Devanagari letter au -> Devanagari letter a + Devanagari vowel sign au

#### In Bengali script confusable

Input	Output	Display	Comment
0986	0985 09BE	আ ; আ	Bengali letter aa -> Bengali letter a + Bengali vowel sign aa
09EA *	0038	৪ ; ৪	Bengali digit four -> Digit eight
09ED *	0039	৭ ; ৭	Bengali digit seven -> Digit nine

#### In Gurmukhi script confusable

Input	Output	Display	Comment
-------	--------	---------	---------

0A06	0A05 0A3E	ਆ ; ਅਾ	Gurmukhi letter aa -> Gurmukhi letter a + Gurmukhi vowel sign aa
0A67 *	0039	ੴ ; 9	Gurmukhi digit one -> Digit nine
0A6A	0038	ੳ ; 8	Gurmukhi digit four -> Digit eight

#### In Gujarati script confusable

Input	Output	Display	Comment
0A86	0A85 0ABE	અ ; અા	Gujarati letter aa -> Gujarati letter a + Gujarati vowel sign aa
0A8D	0A85 0AC5	એ ; એે	Gujarati letter candra e -> Gujarati letter a + Gujarati vowel sign candra e
0A8F	0A85 0AC7	એ ; એે	Gujarati letter e -> Gujarati letter a + Gujarati vowel sign e
0A90	0A85 0AC8	એ ; એૈ	Gujarati letter ai -> Gujarati letter a + Gujarati vowel sign ai
0A91	0A85 0AC9	ઓ ; ઓો	Gujarati letter candra o -> Gujarati letter a + Gujarati vowel sign candra o
0A93	0A85 0ACB	ઓ ; ઓો	Gujarati letter o -> Gujarati letter a + Gujarati vowel sign o
0A94	0A85 0ACC	ઔ ; ઔૌ	Gujarati letter au -> Gujarati letter a + Gujarati vowel sign au
0AE6	0030	૦ ; ૦	Gujarati digit zero -> Digit zero (some fonts have a glyph closer to 006F Latin small letter o)

#### In Oriya script confusable

Input	Output	Display	Comment
0B03 *	0038	୪ ; ୪	Oriya sign visarga -> Digit eight
0B06	0B05 0B3E	ଅା ; ଅା	Oriya letter aa -> Oriya letter a + Oriya vowel sign aa
0B66 *	0030	୦ ; ୦	Oriya digit zero -> Digit zero
0B68 *	0039	୨ ; ୨	Oriya digit two -> Digit 9

#### In Tamil script confusable

Input	Output	Display	Comment
0BD7	0BB3	ௌ ; ள	Tamil au length mark -> Tamil letter lla
0BE7	0B95	௧ ; ௧	Tamil digit one -> Tamil letter ka
0BEE	0B85	௮ ; ௮	Tamil digit eight -> Tamil letter a

#### In Telugu script confusable

Input	Output	Display	Comment
0C02 *	0C66	౦ ; ౦	Telugu sign anusvara -> Telugu digit zero (also confusable with 006F Latin small letter o)

### In Kannada script confusable

Input	Output	Display	Comment
0C82	0CE6	◌̣o ; o	Telugu sign anusvara -> Kannada digit zero (also confusable with 006F Latin small letter o)

### In Cherokee script confusable

Input	Output	Display	Comment
13B6	13C0	G ; G	Cherokee letter lo -> Cherokee letter nah (also confusable with 0047 Latin capital letter g)
13BD	13A9	y ; y	Cherokee letter mu -> Cherokee letter gi (also confusable with 0059 Latin capital letter y)
13CE	0034	4 ; 4	Cherokee letter se -> Digit four
13D2	13A1	R ; R	Cherokee letter sv -> Cherokee letter e (also confusable with 0052 Latin capital letter r)
13D4	13B3	W ; W	Cherokee letter ta -> Cherokee letter la (also confusable with 0057 Latin capital letter w)
13D5	13DA	S ; S	Cherokee letter de -> Cherokee letter du (also confusable with 0053 Latin capital letter s)
13DD	13DE	L ; L	Cherokee letter tla -> Cherokee letter tle (also confusable with 004C Latin capital letter l)
13E3	13C0	G ; G	Cherokee letter tsa -> Cherokee letter nah (also confusable with 0047 Latin capital letter g)
13E4	13D9	V ; V	Cherokee letter tse -> Cherokee letter do (also confusable with 0056 Latin capital letter v)
13E9	13C0	G ; G	Cherokee letter wa -> Cherokee letter nah (also confusable with 0047 Latin capital letter g)
13EE	0036	6 ; 6	Cherokee letter wv -> Digit six
13F3	13C0	G ; G	Cherokee letter yu -> Cherokee letter nah (also confusable with 0047 Latin capital letter g)

### In Canadian Aboriginal Syllabics script confusable

Input	Output	Display	Comment
140C	1427 1401	◌̣ ; ◌̣	Canadian syllabics we -> Canadian syllabics final middle dot + Canadian syllabics e
140D	1401 1427	◌̣̣ ; ◌̣̣	Canadian syllabics west-cree we -> Canadian syllabics e + Canadian syllabics final middle dot
140E	1427 1403	◌̣̣ ; ◌̣̣	Canadian syllabics wi -> Canadian syllabics final middle dot + Canadian syllabics i
140F	1403 1427	◌̣̣̣ ; ◌̣̣̣	Canadian syllabics west-cree wi -> Canadian syllabics I + Canadian syllabics final middle dot
1410	1427 1404	◌̣̣̣ ; ◌̣̣̣	Canadian syllabics wii -> Canadian syllabics final middle dot + Canadian syllabics ii
1411	1404 1427	◌̣̣̣̣ ; ◌̣̣̣̣	Canadian syllabics west-cree wii -> Canadian syllabics ii + Canadian syllabics final middle dot
1412	1427 1405	◌̣̣̣̣ ; ◌̣̣̣̣	Canadian syllabics wo -> Canadian syllabics final middle dot + Canadian syllabics o



1413	1405 1427	ᑭ ; ᑭ	Canadian syllabics west-cree wo -> Canadian syllabics o + Canadian syllabics final middle dot
1414	1427 1406	ᑭ ; ᑭ	Canadian syllabics woo -> Canadian syllabics final middle dot + Canadian syllabics oo
1415	1406 1427	ᑭ ; ᑭ	Canadian syllabics west-cree woo -> Canadian syllabics oo + Canadian syllabics final middle dot
1417	1427 140A	ᑭ ; ᑭ	Canadian syllabics wa -> Canadian syllabics final middle dot + Canadian syllabics a
1418	140A 1427	ᑭ ; ᑭ	Canadian syllabics west-cree wa -> Canadian syllabics a + Canadian syllabics final middle dot
1419	1427 140B	ᑭ ; ᑭ	Canadian syllabics waa -> Canadian syllabics final middle dot + Canadian syllabics aa
141A	140B 1427	ᑭ ; ᑭ	Canadian syllabics west-cree waa -> Canadian syllabics aa + Canadian syllabics final middle dot
1440	1427 1433	ᑭ ; ᑭ	Canadian syllabics pwo -> Canadian syllabics final middle dot + Canadian syllabics po
1441	1433 1427	ᑭ ; ᑭ	Canadian syllabics west-cree pwo -> Canadian syllabics po + Canadian syllabics final middle dot
1442	1427 1434	ᑭ ; ᑭ	Canadian syllabics pwoo -> Canadian syllabics final middle dot + Canadian syllabics poo
1443	1434 1427	ᑭ ; ᑭ	Canadian syllabics west-cree pwoo -> Canadian syllabics poo + Canadian syllabics final middle dot
1444	1427 1438	ᑭ ; ᑭ	Canadian syllabics pwa -> Canadian syllabics final middle dot + Canadian syllabics pa
1445	1438 1427	ᑭ ; ᑭ	Canadian syllabics west-cree pwa -> Canadian syllabics pa + Canadian syllabics final middle dot
1446	1427 1439	ᑭ ; ᑭ	Canadian syllabics pwaa -> Canadian syllabics final middle dot + Canadian syllabics paa
1447	1439 1427	ᑭ ; ᑭ	Canadian syllabics west-cree pwaa -> Canadian syllabics paa + Canadian syllabics final middle dot
1457	1427 144C	ᑭ ; ᑭ	Canadian syllabics twe -> Canadian syllabics final middle dot + Canadian syllabics te
1458	144C 1427	ᑭ ; ᑭ	Canadian syllabics west-cree twe -> Canadian syllabics te + Canadian syllabics final middle dot
1459	1427 144E	ᑭ ; ᑭ	Canadian syllabics twi -> Canadian syllabics final middle dot + Canadian syllabics ti
145A	144E 1427	ᑭ ; ᑭ	Canadian syllabics west-cree twi -> Canadian syllabics ti + Canadian syllabics final middle dot
145B	1427 144F	ᑭ ; ᑭ	Canadian syllabics twii -> Canadian syllabics final middle dot + Canadian syllabics tii
145C	144F 1427	ᑭ ; ᑭ	Canadian syllabics west-cree twii -> Canadian syllabics tii + Canadian syllabics final middle dot
145D	1427 1450	ᑭ ; ᑭ	Canadian syllabics two -> Canadian syllabics final middle dot + Canadian syllabics to
145E	1450 1427	ᑭ ; ᑭ	Canadian syllabics west-cree two -> Canadian syllabics to + Canadian syllabics final middle dot
145F	1427 1451	ᑭ ; ᑭ	Canadian syllabics twoo -> Canadian syllabics final middle dot + Canadian syllabics too
1460	1451 1427	ᑭ ; ᑭ	Canadian syllabics west-cree twoo -> Canadian syllabics too + Canadian syllabics final middle dot

1461	1427 1455	·C ; ·C	Canadian syllabics twa -> Canadian syllabics final middle dot + Canadian syllabics ta
1462	1455 1427	C· ; C·	Canadian syllabics west-cree twa -> Canadian syllabics ta + Canadian syllabics final middle dot
1463	1427 1456	·Ć ; ·Ć	Canadian syllabics twaa -> Canadian syllabics final middle dot + Canadian syllabics taa
1464	1456 1427	Ć· ; Ć·	Canadian syllabics west-cree twaa -> Canadian syllabics taa + Canadian syllabics final middle dot
1467	144C 144A	U' ; U'	Canadian syllabics tte -> Canadian syllabics te + Canadian syllabics west-cree p
1468	144E 144A	Ń' ; Ń'	Canadian syllabics tti -> Canadian syllabics ti + Canadian syllabics west-cree p
1469	1450 144A	ᑄ' ; ᑄ'	Canadian syllabics tto -> Canadian syllabics to + Canadian syllabics west-cree p
146A	1455 144A	C' ; C'	Canadian syllabics tta -> Canadian syllabics ta + Canadian syllabics west-cree p
1474	1427 146B	·q ; ·q	Canadian syllabics kwe -> Canadian syllabics final middle dot + Canadian syllabics ke
1475	146B 1427	q· ; q·	Canadian syllabics west-cree kwe -> Canadian syllabics ke + Canadian syllabics final middle dot
1476	1427 146D	·p ; ·p	Canadian syllabics kwi -> Canadian syllabics final middle dot + Canadian syllabics ki
1477	146D 1427	p· ; p·	Canadian syllabics west-cree kwi -> Canadian syllabics ki + Canadian syllabics final middle dot
1478	1427 146E	·ᑭ ; ·ᑭ	Canadian syllabics kwii -> Canadian syllabics final middle dot + Canadian syllabics kii
1479	146E 1427	ᑭ· ; ᑭ·	Canadian syllabics west-cree kwii -> Canadian syllabics kii + Canadian syllabics final middle dot
147A	1427 146F	·d ; ·d	Canadian syllabics kwo -> Canadian syllabics final middle dot + Canadian syllabics ko
147B	146F 1427	d· ; d·	Canadian syllabics west-cree kwo -> Canadian syllabics ko + Canadian syllabics final middle dot
147C	1427 1470	·ᑭ ; ·ᑭ	Canadian syllabics kwoo -> Canadian syllabics final middle dot + Canadian syllabics koo
147D	1470 1427	ᑭ· ; ᑭ·	Canadian syllabics west-cree kwoo -> Canadian syllabics koo + Canadian syllabics final middle dot
147E	1427 1472	·b ; ·b	Canadian syllabics kwa -> Canadian syllabics final middle dot + Canadian syllabics ka
147F	1472 1427	b· ; b·	Canadian syllabics west-cree kwa -> Canadian syllabics ka + Canadian syllabics final middle dot
1480	1427 1473	·ᑭ ; ·ᑭ	Canadian syllabics kwaa -> Canadian syllabics final middle dot + Canadian syllabics kaa
1481	1473 1427	ᑭ· ; ᑭ·	Canadian syllabics west-cree kwaa -> Canadian syllabics kaa + Canadian syllabics final middle dot
1485	146B 144A	q' ; q'	Canadian syllabics south-slavey keh -> Canadian syllabics ke + Canadian syllabics west-cree p
1486	146D 144A	p' ; p'	Canadian syllabics south-slavey kih -> Canadian syllabics ki + Canadian syllabics west-cree p
1487	146F 144A	d' ; d'	Canadian syllabics south-slavey koh -> Canadian syllabics ko + Canadian syllabics west-cree p

1488	1472 144A	ᑲ' ; ᑲ'	Canadian syllabics south-slavey kah -> Canadian syllabics ka + Canadian syllabics west-cree p
1492	1427 1489	ᑲ̣ ; ᑲ̣	Canadian syllabics cwe -> Canadian syllabics final middle dot + Canadian syllabics ce
1493	1489 1427	ᑲ̣̣ ; ᑲ̣̣	Canadian syllabics west-cree cwe -> Canadian syllabics ce + Canadian syllabics final middle dot
1494	1427 148B	ᑲ̣̣̣ ; ᑲ̣̣̣	Canadian syllabics cwi -> Canadian syllabics final middle dot + Canadian syllabics ci
1495	148B 1427	ᑲ̣̣̣̣ ; ᑲ̣̣̣̣	Canadian syllabics west-cree cwi -> Canadian syllabics ci + Canadian syllabics final middle dot
1496	1427 148C	ᑲ̣̣̣̣̣ ; ᑲ̣̣̣̣̣	Canadian syllabics cwii -> Canadian syllabics final middle dot + Canadian syllabics cii
1497	148C 1427	ᑲ̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣	Canadian syllabics west-cree cwii -> Canadian syllabics cii + Canadian syllabics final middle dot
1498	1427 148D	ᑲ̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣	Canadian syllabics cwo -> Canadian syllabics final middle dot + Canadian syllabics co
1499	148D 1427	ᑲ̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣	Canadian syllabics west-cree cwo -> Canadian syllabics co + Canadian syllabics final middle dot
149A	1427 148E	ᑲ̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣	Canadian syllabics cwoo -> Canadian syllabics final middle dot + Canadian syllabics coo
149B	148E 1427	ᑲ̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree cwoo -> Canadian syllabics coo + Canadian syllabics final middle dot
149C	1427 1490	ᑲ̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics cwa -> Canadian syllabics final middle dot + Canadian syllabics ca
149D	1490 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree cwa -> Canadian syllabics ca + Canadian syllabics final middle dot
149E	1427 1491	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics cwaā -> Canadian syllabics final middle dot + Canadian syllabics caa
149F	1491 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree cwaā -> Canadian syllabics caa + Canadian syllabics final middle dot
14AC	1427 14A3	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics mwe -> Canadian syllabics final middle dot + Canadian syllabics me
14AD	14A3 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree mwe -> Canadian syllabics me + Canadian syllabics final middle dot
14AE	1427 14A5	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics mwi -> Canadian syllabics final middle dot + Canadian syllabics mi
14AF	14A5 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree mwi -> Canadian syllabics mi + Canadian syllabics final middle dot
14B0	1427 14A6	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics mwii -> Canadian syllabics final middle dot + Canadian syllabics mii
14B1	14A6 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree mwii -> Canadian syllabics mii + Canadian syllabics final middle dot
14B2	1427 14A7	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics mwo -> Canadian syllabics final middle dot + Canadian syllabics mo
14B3	14A7 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree mwo -> Canadian syllabics mo + Canadian syllabics final middle dot
14B4	1427 14A8	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics mwoo -> Canadian syllabics final middle dot + Canadian syllabics moo
14B5	14A8 1427	ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣ ; ᑲ̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣̣	Canadian syllabics west-cree mwoo -> Canadian syllabics moo + Canadian syllabics final middle dot

14B6	1427 14AA	·L ; ·L	Canadian syllabics mwa -> Canadian syllabics final middle dot + Canadian syllabics ma
14B7	14AA 1427	L· ; L·	Canadian syllabics west-cree mwa -> Canadian syllabics ma + Canadian syllabics final middle dot
14B8	1427 14AB	·Ḷ ; ·Ḷ	Canadian syllabics mwaa -> Canadian syllabics final middle dot + Canadian syllabics maa
14B9	14AB 1427	Ḷ· ; Ḷ·	Canadian syllabics west-cree mwaa -> Canadian syllabics maa + Canadian syllabics final middle dot
14BE	0032	ᓂ ; 2	Canadian syllabics athapascan m
14BF	0032	ᓂ ; 2	Canadian syllabics sayisi m -> Digit two
14C9	1427 14C0	·ᓄ ; ·ᓄ	Canadian syllabics I -> Canadian syllabics final middle dot + Canadian syllabics ne
14CA	14C0 1427	ᓄ· ; ᓄ·	Canadian syllabics west-cree I -> Canadian syllabics ne + Canadian syllabics final middle dot
14CB	1427 14C7	·ᓇ ; ·ᓇ	Canadian syllabics nwa -> Canadian syllabics final middle dot + Canadian syllabics na
14CC	14C7 1427	ᓇ· ; ᓇ·	Canadian syllabics west-cree nwa -> Canadian syllabics na + Canadian syllabics final middle dot
14CD	1427 14C8	·ᓈ ; ·ᓈ	Canadian syllabics nwaa -> Canadian syllabics final middle dot + Canadian syllabics naa
14CE	14C8 14AB	ᓈ· ; ᓈ·	Canadian syllabics west-cree nwaa -> Canadian syllabics naa + Canadian syllabics final middle dot
14DC	1427 14D3	·ᓌ ; ·ᓌ	Canadian syllabics lwe -> Canadian syllabics final middle dot + Canadian syllabics le
14DD	14D3 1427	ᓌ· ; ᓌ·	Canadian syllabics west-cree lwe -> Canadian syllabics le + Canadian syllabics final middle dot
14DE	1427 14D5	·ᓎ ; ·ᓎ	Canadian syllabics lwi -> Canadian syllabics final middle dot + Canadian syllabics li
14DF	14D5 1427	ᓎ· ; ᓎ·	Canadian syllabics west-cree lwi -> Canadian syllabics li + Canadian syllabics final middle dot
14E0	1427 14D6	·ᓐ ; ·ᓐ	Canadian syllabics lwii -> Canadian syllabics final middle dot + Canadian syllabics lii
14E1	14D6 1427	ᓐ· ; ᓐ·	Canadian syllabics west-cree lwii -> Canadian syllabics lii + Canadian syllabics final middle dot
14E2	1427 14D7	·ᓑ ; ·ᓑ	Canadian syllabics lwo -> Canadian syllabics final middle dot + Canadian syllabics lo
14E3	14D7 1427	ᓑ· ; ᓑ·	Canadian syllabics west-cree lwo -> Canadian syllabics lo + Canadian syllabics final middle dot
14E4	1427 14D8	·ᓓ ; ·ᓓ	Canadian syllabics lwoo -> Canadian syllabics final middle dot + Canadian syllabics loo
14E5	14D8 1427	ᓓ· ; ᓓ·	Canadian syllabics west-cree lwoo -> Canadian syllabics loo + Canadian syllabics final middle dot
14E6	1427 14DA	·ᓔ ; ·ᓔ	Canadian syllabics lwa -> Canadian syllabics final middle dot + Canadian syllabics la
14E7	14DA 1427	ᓔ· ; ᓔ·	Canadian syllabics west-cree lwa -> Canadian syllabics la + Canadian syllabics final middle dot
14E8	1427 14DB	·ᓖ ; ·ᓖ	Canadian syllabics lwaa -> Canadian syllabics final middle dot + Canadian syllabics laa

14E9	14DB 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree lwaa -> Canadian syllabics laa + Canadian syllabics final middle dot
14F6	1427 14ED	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swe -> Canadian syllabics final middle dot + Canadian syllabics se
14F7	14ED 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swe -> Canadian syllabics se + Canadian syllabics final middle dot
14F8	1427 14EF	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swi -> Canadian syllabics final middle dot + Canadian syllabics si
14F9	14EF 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swi -> Canadian syllabics si + Canadian syllabics final middle dot
14FA	1427 14F0	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swii -> Canadian syllabics final middle dot + Canadian syllabics sii
14FB	14F0 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swii -> Canadian syllabics sii + Canadian syllabics final middle dot
14FC	1427 14F1	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swo -> Canadian syllabics final middle dot + Canadian syllabics so
14FD	14F1 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swo -> Canadian syllabics so + Canadian syllabics final middle dot
14FE	1427 14F2	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swoo -> Canadian syllabics final middle dot + Canadian syllabics soo
14FF	14F2 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swoo -> Canadian syllabics soo + Canadian syllabics final middle dot
1500	1427 14F4	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swa -> Canadian syllabics final middle dot + Canadian syllabics sa
1501	14F4 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swa -> Canadian syllabics sa + Canadian syllabics final middle dot
1502	1427 14F5	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics swaa -> Canadian syllabics final middle dot + Canadian syllabics saa
1503	14F5 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree swaa -> Canadian syllabics saa + Canadian syllabics final middle dot
150C	150B 1438	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics unavut scwa -> Canadian syllabics Naskapi s-w + Canadian syllabics pa
150D	150B 1455	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics unavut stwa -> Canadian syllabics Naskapi s-w + Canadian syllabics ta
150E	150B 1472	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics unavut skwa -> Canadian syllabics Naskapi s-w + Canadian syllabics ka
150F	150B 1490	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics unavut scwa -> Canadian syllabics Naskapi s-w + Canadian syllabics ca
1517	1427 1510	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics shwe -> Canadian syllabics final middle dot + Canadian syllabics she
1518	1510 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree shwe -> Canadian syllabics she + Canadian syllabics final middle dot
1519	1427 1511	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics shwi -> Canadian syllabics final middle dot + Canadian syllabics shi
151A	1511 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree shwi -> Canadian syllabics shi + Canadian syllabics final middle dot
151B	1427 1512	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics shwii -> Canadian syllabics final middle dot + Canadian syllabics shii
151C	1512 1427	ᑭᑦ ; ᑭᑦᑦ	Canadian syllabics west-cree shwii -> Canadian syllabics shii + Canadian syllabics final middle dot

151D	1427 1513	ᑕ ; ᑕ	Canadian syllabics una -> Canadian syllabics final middle dot + Canadian syllabics sho
151E	1513 1427	ᑕ ; ᑕ	Canadian syllabics west-cree una -> Canadian syllabics sho + Canadian syllabics final middle dot
151F	1427 1514	ᑕ ; ᑕ	Canadian syllabics shwoo -> Canadian syllabics final middle dot + Canadian syllabics shoo
1520	1514 1427	ᑕ ; ᑕ	Canadian syllabics west-cree shwoo -> Canadian syllabics shoo + Canadian syllabics final middle dot
1521	1427 1515	ᑕ ; ᑕ	Canadian syllabics shwa -> Canadian syllabics final middle dot + Canadian syllabics sha
1522	1515 1427	ᑕ ; ᑕ	Canadian syllabics west-cree shwa -> Canadian syllabics sha + Canadian syllabics final middle dot
1523	1427 1516	ᑕ ; ᑕ	Canadian syllabics shwaa -> Canadian syllabics final middle dot + Canadian syllabics shaa
1524	1516 1427	ᑕ ; ᑕ	Canadian syllabics west-cree shwaa -> Canadian syllabics shaa + Canadian syllabics final middle dot
1526	0034	4 ; 4	Canadian syllabics ye -> Digit four
152F	1427 1526	ᑕ ; ᑕ	Canadian syllabics ywe -> Canadian syllabics final middle dot + Canadian syllabics ye
1530	1526 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywe -> Canadian syllabics ye + Canadian syllabics final middle dot
1531	1427 1528	ᑕ ; ᑕ	Canadian syllabics ywi -> Canadian syllabics final middle dot + Canadian syllabics yi
1532	1528 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywi -> Canadian syllabics yi + Canadian syllabics final middle dot
1533	1427 1529	ᑕ ; ᑕ	Canadian syllabics ywii -> Canadian syllabics final middle dot + Canadian syllabics yii
1534	1529 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywii -> Canadian syllabics yii + Canadian syllabics final middle dot
1535	1427 152A	ᑕ ; ᑕ	Canadian syllabics ywo -> Canadian syllabics final middle dot + Canadian syllabics yo
1536	152A 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywo -> Canadian syllabics yo + Canadian syllabics final middle dot
1537	1427 152B	ᑕ ; ᑕ	Canadian syllabics ywoo -> Canadian syllabics final middle dot + Canadian syllabics yoo
1538	152B 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywoo -> Canadian syllabics yoo + Canadian syllabics final middle dot
1539	1427 152D	ᑕ ; ᑕ	Canadian syllabics ywa -> Canadian syllabics final middle dot + Canadian syllabics ya
153A	152D 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywa -> Canadian syllabics ya + Canadian syllabics final middle dot
153B	1427 152E	ᑕ ; ᑕ	Canadian syllabics ywaa -> Canadian syllabics final middle dot + Canadian syllabics yaa
153C	152E 1427	ᑕ ; ᑕ	Canadian syllabics west-cree ywaa -> Canadian syllabics yaa + Canadian syllabics final middle dot
154E	1427 154C	ᑕ ; ᑕ	Canadian syllabics rwaa -> Canadian syllabics final middle dot + Canadian syllabics raa
154F	154C 1427	ᑕ ; ᑕ	Canadian syllabics west-cree rwaa -> Canadian syllabics raa + Canadian syllabics final middle dot

155B	1427 155A	ᑦᑦ ; ᑦᑦ	Canadian syllabics fwaa -> Canadian syllabics final middle dot + Canadian syllabics faa
155C	155A 1427	ᑦᑦ ; ᑦᑦ	Canadian syllabics west-cree fwaa -> Canadian syllabics faa + Canadian syllabics final middle dot
1568	1427 1567	ᑦᑦ ; ᑦᑦ	Canadian syllabics thwaa -> Canadian syllabics final middle dot + Canadian syllabics thaa
1569	1567 1427	ᑦᑦ ; ᑦᑦ	Canadian syllabics west-cree thwaa -> Canadian syllabics thaa + Canadian syllabics final middle dot
156F	002A	* ; *	Canadian syllabics tth -> Asterisk (one or two?)
157D	1541	x ; x	Canadian syllabics hk -> Canadian syllabics sayisi yi (also confusable with 0078 Latin small letter x)
157E	1550 146C	ᑦᑦ ; ᑦᑦ	Canadian syllabics qaai -> Canadian syllabics r + Canadian syllabics kaai
157F	1550 146D	ᑦᑦ ; ᑦᑦ	Canadian syllabics qi -> Canadian syllabics r + Canadian syllabics ki
1580	1550 146E	ᑦᑦ ; ᑦᑦ	Canadian syllabics qii -> Canadian syllabics r + Canadian syllabics kii
1581	1550 146F	ᑦᑦ ; ᑦᑦ	Canadian syllabics qo -> Canadian syllabics r + Canadian syllabics ko
1582	1550 1470	ᑦᑦ ; ᑦᑦ	Canadian syllabics qoo -> Canadian syllabics r + Canadian syllabics koo
1583	1550 1472	ᑦᑦ ; ᑦᑦ	Canadian syllabics qa -> Canadian syllabics r + Canadian syllabics ka
1584	1550 1473	ᑦᑦ ; ᑦᑦ	Canadian syllabics qaa -> Canadian syllabics r + Canadian syllabics kaa
1585	1550 1483	ᑦᑦ ; ᑦᑦ	Canadian syllabics q -> Canadian syllabics r + Canadian syllabics k
158E	1595 148A	ᑦᑦ ; ᑦᑦ	Canadian syllabics ngaai -> Canadian syllabics ng + Canadian syllabics caai
158F	1595 148B	ᑦᑦ ; ᑦᑦ	Canadian syllabics ngi -> Canadian syllabics ng + Canadian syllabics ci
1590	1595 148C	ᑦᑦ ; ᑦᑦ	Canadian syllabics ngii -> Canadian syllabics ng + Canadian syllabics cii
1591	1595 148D	ᑦᑦ ; ᑦᑦ	Canadian syllabics ngo -> Canadian syllabics ng + Canadian syllabics co
1592	1595 148E	ᑦᑦ ; ᑦᑦ	Canadian syllabics ngoo -> Canadian syllabics ng + Canadian syllabics coo
1593	1595 1490	ᑦᑦ ; ᑦᑦ	Canadian syllabics nga -> Canadian syllabics ng + Canadian syllabics ca
1594	1595 1491	ᑦᑦ ; ᑦᑦ	Canadian syllabics ngaa -> Canadian syllabics ng + Canadian syllabics caa
15C4	2200	∇ ; ∇	Canadian syllabics carrier ghu -> For all
15EA	15DE	D ; D	Canadian syllabics carrier pe -> Canadian syllabics carrier the (also confusable with 0044 Latin capital letter d)
15F8	15F7	B ; B	Canadian syllabics carrier kee -> Canadian syllabics khe (also confusable with 0042 Latin capital letter b)
15F9	15F7	B ; B	Canadian syllabics carrier khi -> Canadian syllabics khe (also confusable with 0042 Latin capital letter b)

1602	1490	ᓆ ; ᓆ	Canadian syllabics carrier nu -> Canadian syllabics ca
1603	1489	ᓇ ; ᓇ	Canadian syllabics carrier no -> Canadian syllabics ce
1604	14D3	ᓈ ; ᓈ	Canadian syllabics carrier ne -> Canadian syllabics le
1607	14DA	ᓉ ; ᓉ	Canadian syllabics carrier na -> Canadian syllabics la
1616	0032	ᓊ ; 2	Canadian syllabics carrier jo -> Digit two (many duplicates)
1622	1543	ᓋ ; ᓋ	Canadian syllabics carrier lu -> Canadian syllabics r-cree re
1623	1546	ᓌ ; ᓌ	Canadian syllabics carrier lo -> Canadian syllabics r-cree ri
1624	154A	ᓍ ; ᓍ	Canadian syllabics carrier le -> Canadian syllabics west-cree lo
1624	154A	ᓍ ; ᓍ	Canadian syllabics carrier le -> Canadian syllabics west-cree lo
162E	2127	ᓏ ; ᓏ	Canadian syllabics carrier lhu -> inverted ohm sign
162F	2126	ᓐ ; ᓐ	Canadian syllabics carrier lho -> ohm sign
1634	162E	ᓑ ; ᓑ	Canadian syllabics carrier tlhu -> Canadian syllabics carrier lhu
1635	162F	ᓒ ; ᓐ	Canadian syllabics carrier tlho -> Canadian syllabics carrier lho
1656	15F7	ᓔ ; B	Canadian syllabics carrier she -> Canadian syllabics khe (also confusable with 0042 Latin capital letter b)
1657	15F7	ᓕ ; B	Canadian syllabics carrier she -> Canadian syllabics khe (also confusable with 0042 Latin capital letter b)
1658	15F7	ᓆ ; B	Canadian syllabics carrier shi -> Canadian syllabics khe (also confusable with 0042 Latin capital letter b)
166e	1541	ᓗ ; x	Canadian syllabics full stop -> Canadian syllabics sayisi yi (also confusable with 0078 Latin small letter x)
166F	1550 146B	ᓘᓐ ; ᓘᓐ	Canadian syllabics qai -> Canadian syllabics r + Canadian syllabics ke
1670	1595 1489	ᓘᓇ ; ᓘᓇ	Canadian syllabics ngai -> Canadian syllabics ng + Canadian syllabics ce
1671	1596 148B	ᓘᓇᓐ ; ᓘᓇᓐ	Canadian syllabics nngi -> Canadian syllabics nng + Canadian syllabics ci
1672	1596 148C	ᓘᓇᓐᓐ ; ᓘᓇᓐᓐ	Canadian syllabics nngii -> Canadian syllabics nng + Canadian syllabics cii
1673	1596 148D	ᓘᓇᓐᓐᓐ ; ᓘᓇᓐᓐᓐ	Canadian syllabics nngo -> Canadian syllabics nng + Canadian syllabics co
1674	1596 148E	ᓘᓇᓐᓐᓐᓐ ; ᓘᓇᓐᓐᓐᓐ	Canadian syllabics nngoo -> Canadian syllabics nng + Canadian syllabics coo
1675	1596 1490	ᓘᓇᓐᓐᓐᓐᓐ ; ᓘᓇᓐᓐᓐᓐᓐ	Canadian syllabics nnga -> Canadian syllabics nng + Canadian syllabics ca
1676	1596 1491	ᓘᓇᓐᓐᓐᓐᓐᓐ ; ᓘᓇᓐᓐᓐᓐᓐᓐ	Canadian syllabics nngaa -> Canadian syllabics nng + Canadian syllabics caa



**In mixed script (Greek-Latin) confusable, lower case**

Input	Output	Display	Comment
0326	0327	ⵛ ; ⵛ	Combining comma below – Combining cedilla
0342	0303	ⵀ ; ⵀ	Combining greek perispomeni -> Combining tilde
03B1 *	0061	α ; a	Greek small letter alpha -> Latin small letter alpha
03B3	0263	γ ; γ	Greek small letter gamma -> Latin small letter gamma
03B5 *	025B	ε ; ε	Greek small letter epsilon -> Latin small letter open e
03B7	014B	η ; η	Greek small letter eta -> Latin small letter eng
03B7 *	006E 0329	η̣ ; η̣	Greek small letter eta -> Latin small letter n + Combining vertical line below
03B9 *	0131	ι̣ ; ι̣	Greek small letter iota -> Latin small letter dotless i (TR36 says 0069)
03BA *	0138	κ ; κ	Greek small letter kappa -> Latin small letter kra
03BD *	0076	ν ; ν	Greek small letter nu -> Latin small letter v
03BF *	006F	ο ; ο	Greek small letter omicron -> Latin small letter o
03C0	006E	π ; n	Greek small letter pi -> Latin small letter n
03C1 *	0070	ρ ; ρ	Greek small letter rho -> Latin small letter p
03C5 *	028B	υ ; υ	Greek small letter upsilon -> Latin small letter v with hook
03C6	0278	φ ; φ	Greek small letter phi -> Latin small letter phi (closer to 03d5 Greek phi symbol φ itself IDN-mapped to 03c6)
03C7 *	0078	χ ; x	Greek small letter chi -> Latin small letter x
03E9	01A8	ⲉ ; e	Coptic small letter hori -> Latin small letter tone two
03F2	0063	σ ; c	Greek lunate sigma symbol -> Latin small letter c (not needed as 03F2 is IDN-mapped to 03C3 σ not confusable)
03F3 *	006A	Ϸ ; j	Greek letter yot -> Latin small letter j

**In mixed script (Cyrillic-Latin) confusable, lower case**

Input	Output	Display	Comment
0430 *	0061	а ; a	Cyrillic small letter a -> Latin small letter a
0432 *	0299	в ; B	Cyrillic small letter ve -> Latin small capital b
0433 *	0072	г ; r	Cyrillic small letter ghe -> Latin small letter r

0435 *	0065	е ; е	Cyrillic small letter ie -> Latin small letter e
0437	025c	э ; э	Cyrillic small letter ze -> Latin small letter reversed open e
0437 *	025C	э ; э	Cyrillic small letter ze -> Latin small letter reversed open e
043A *	0138	к ; к	Cyrillic small letter ka -> Latin small letter kra
043D *	029C	н ; н	Cyrillic small letter en -> Latin letter small capital h
043E *	006F	о ; о	Cyrillic small letter o -> Latin small letter o
0440 *	0070	р ; р	Cyrillic small letter er -> Latin small letter p
0441 *	0063	с ; с	Cyrillic small letter es -> Latin small letter c
0443 *	0079	у ; у	Cyrillic small letter u -> Latin small letter y
0445 *	0078	х ; х	Cyrillic small letter ha -> Latin small letter x
0447	0265	ч ; ч	Latin Cyrillic che -> Latin small letter turned h
044A	0062	ь ; ь	Cyrillic small letter soft sign -> Latin small letter b
0455 *	0073	ѕ ; ѕ	Cyrillic small letter dze -> Latin small letter s
0456 *	0069	і ; і	Cyrillic small letter Byelorussian-ukrainian I -> Latin small letter i
0458 *	006A	ј ; ј	Cyrillic small letter je -> Latin small letter j
0475 *	0076	ѵ ; ѵ	Cyrillic small letter izhitsa -> Latin small letter v
048D	0180	ѣ ; ѣ	Cyrillic small letter semisoft sign -> Latin small letter b with stroke
04BB *	0068	һ ; һ	Cyrillic small letter shha -> Latin small letter h
04D5 *	0061 0065	æ ; æ	Cyrillic small ligature a ie -> Latin small letter a + Latin small letter e (TR36 says 00E6)
04D9 *	0259	ə ; ə	Cyrillic small letter schwa -> Latin small letter schwa
04E1	0292	ӡ ; ӡ	Cyrillic small letter abkhasian dze -> Latin small letter ezh
04E1 *	01BA 0321	ӡ ; ӡ	Cyrillic small letter abkhasian dze -> Latin small letter ezh with tail + Combining palatalized hook below (????)
04E9 *	0275	ө ; ө	Cyrillic small letter barred o -> Latin small letter barred o
0501 *	0064	Ꞥ ; Ꞥ	Cyrillic small letter komi de -> Latin small letter d
050D	0262	Ꞥ ; Ꞥ	Cyrillic small letter komi sje -> Latin letter small capital g

**In mixed script (Cyrillic-Greek) confusable, lower case**

Input	Output	Display	Comment
043F *	03C0	п ; π	Cyrillic small letter pe -> Greek small letter pi (also confusable with 006E Latin small letter n)
0442 *	03C4	т ; τ	Cyrillic small letter te -> Greek small letter tau
0454 *	03B5	е ; ε	Cyrillic small letter krainian ie -> Greek small letter epsilon

**In mixed script (Armenian-Latin) confusable, lower case**

Input	Output	Display	Comment
0272	0580	ղ ; ꞥ	Latin small letter n with left hook -> Armenian small letter reh
0566	0071	զ ; q	Armenian small letter za -> Latin small letter q
056A	0063	ժ ; d	Armenian small letter zhe -> Latin small letter d
0570 *	0068	հ ; h	Armenian small letter ho -> Latin small letter h
0572	014B	ղ ; ŋ	Armenian small letter ghad -> Latin small letter eng (closer to 0273 Latin small letter n with retroflex hook)
0575	006A	յ ; j	Armenian small letter yi -> Latin small letter j
0578 *	006E	ն ; n	Armenian small letter vo -> Latin small letter n
057D *	0075	ւ ; u	Armenian small letter I -> Latin small letter u
0581 *	0067	ց ; g	Armenian small letter co -> Latin small letter g
0584 *	0066	ֆ ; f	Armenian small letter keh -> Latin small letter f
0585 *	006F	օ ; o	Armenian small letter oh -> Latin small letter o

**In mixed script (Armenian-Greek) confusable, lower case**

Input	Output	Display	Comment
056E	03B4	ժ ; δ	Armenian small letter ca -> Greek small letter delta

**In mixed script (Bengali-Latin) confusable**

Input	Output	Display	Comment
09E6 *	006F	০ ; o	Bengali digit zero -> Latin small letter o

**In mixed script (Gurmukhi-Latin) confusable**

Input	Output	Display	Comment
0A66 *	006F	੦ ; o	Gurmukhi digit zero -> Latin small letter o

**In mixed script (Gujarati-Devanagari) confusable**

Input	Output	Display	Comment
0ABD	093D	ઁ ; s	Gujarati sign avagraha -> Devanagari sign avagraha (Gujarati glyph incorrect in some fonts as above but eventually corrected)

**In mixed script (Telugu-Latin) confusable**

Input	Output	Display	Comment
0C66 *	006F	౦ ; o	Telugu digit zero -> Latin small letter o

**In mixed script (Kannada-Latin) confusable**

Input	Output	Display	Comment
0CE6 *	006F	೦ ; o	Kannada digit zero -> Latin small letter o

**In mixed script (Kannada-Devanagari) confusable**

Input	Output	Display	Comment
0CBD	093D	ॠ ; s	Kannada sign -> avagraha -> Devanagari sign avagraha (character not in Unicode 3.2)

**In mixed script (Thai-Latin) confusable**

Input	Output	Display	Comment
0E50	006F	๐ ; o	Thai digit zero -> Latin small letter o

**In mixed script (Lao-Latin) confusable**

Input	Output	Display	Comment
0ED0	006F	໐ ; o	Thai digit zero -> Latin small letter o

**In mixed script (Georgian-Latin) confusable**

Input	Output	Display	Comment
10E7 *	0079	ყ ; y	Georgian letter qar -> Latin small letter y
10DB	0061	ა ; a	Georgian letter man -> Latin small letter a
10F3	0292	ჟ ; ž	Georgian letter we -> Latin small letter ezh

**In mixed script (Cherokee-Latin) confusable**

Input	Output	Display	Comment
13A0	0044	Ꭰ ; D	Cherokee letter a -> Latin capital letter d
13A1	0052	Ꭱ ; R	Cherokee letter e -> Latin capital letter r
13A2	0054	Ꭲ ; T	Cherokee letter I -> Latin capital letter t
13A5	0069	Ꭶ ; i	Cherokee letter v -> Latin small letter i

13A9	0059	ʏ ; Y	Cherokee letter gi -> Latin capital letter y
13AA	0041	Ꭰ ; A	Cherokee letter go -> Latin capital letter a
13AB	004A	Ꭱ ; J	Cherokee letter gu -> Latin capital letter j
13AC	0045	Ꭲ ; E	Cherokee letter gv -> Latin capital letter e
13B3	0057	Ꭳ ; W	Cherokee letter la -> Latin capital letter w
13B7	004D	Ꭴ ; M	Cherokee letter lu -> Latin capital letter m
13BB	0048	Ꭵ ; H	Cherokee letter mi -> Latin capital letter h
13BE	019F	Ꭶ ; Ø	Cherokee letter na -> Latin capital letter o with middle tilde
13C0	0047	Ꭷ ; G	Cherokee letter nah -> Latin capital letter g
13C2	0068	Ꭸ ; h	Cherokee letter ni -> Latin small letter h
13C3	005A	Ꭹ ; Z	Cherokee letter noi -> Latin capital letter z
13CF	0062	Ꭺ ; b	Cherokee letter si -> Latin small letter b
13D9	0056	Ꭻ ; V	Cherokee letter do -> Latin capital letter v
13DA	0053	Ꭼ ; S	Cherokee letter du -> Latin capital letter s
13DE	004C	Ꭽ ; L	Cherokee letter tle -> Latin capital letter l
13DF	0043	Ꭾ ; C	Cherokee letter tli -> Latin capital letter c
13E2	0050	Ꭿ ; P	Cherokee letter tlv -> Latin capital letter p
13E6	004B	Ꮀ ; K	Cherokee letter tso -> Latin capital letter k
13E7	0044	Ꮁ ; d	Cherokee letter tsu -> Latin small letter d
13F2	0068	Ꮂ ; h	Cherokee letter yo -> Latin small letter h
13F4	0042	Ꮃ ; B	Cherokee letter yv -> Latin capital letter b

**In mixed script (Cherokee-Greek) confusable**

Input	Output	Display	Comment
13B1	0393	Ꭰ ; Γ	Cherokee letter hu -> Greek capital letter gamma

**In mixed script (Canadian Aboriginal Syllabics – Latin) confusable**

Input	Output	Display	Comment
-------	--------	---------	---------

142F	0056	V ; V	Canadian syllabics pe -> Latin capital letter v
144C	0055	U ; U	Canadian syllabics te -> Latin capital letter u
146D	0050	p ; P	Canadian syllabics ki -> Latin capital letter p
146F	0064	d ; d	Canadian syllabics ko -> Latin small letter d
148D	004A	J ; J	Canadian syllabics co -> Latin capital letter j
14AA	004C	L ; L	Canadian syllabics ma -> Latin capital letter v
1515	0053	S ; S	Canadian syllabics sha -> Latin capital letter s
1541	0078	x ; x	Canadian syllabics sayisi yi -> Latin small letter x
157C	0048	H ; H	Canadian syllabics unavut h -> Latin capital letter h
1587	0052	R ; R	Canadian syllabics tlhi -> Latin capital letter r
1589	0281	ʀ ; ʀ	Canadian syllabics tlha -> Latin small capital inverted r
15AF	0062	b ; b	Canadian syllabics aivilik b -> Latin small letter b
15B4	0046	F ; F	Canadian syllabics blackfoot we -> Latin capital letter f
15C5	0041	A ; A	Canadian syllabics carrier gho -> Latin capital letter a
15DE	0044	D ; D	Canadian syllabics carrier the -> Latin capital letter a
15EF	0057	W ; W	Canadian syllabics carrier gu -> Latin capital letter w
15F0	004D	M ; M	Canadian syllabics carrier go -> Latin capital letter m
15F7	0042	B ; B	Canadian syllabics carrier khe -> Latin capital letter b
166D	0058	X ; X	Canadian syllabics chi sign -> Latin capital letter x

**In mixed script (Canadian Aboriginal Syllabics – Greek) confusable**

Input	Output	Display	Comment
1403	0394	Δ ; Δ	Canadian syllabics I -> Greek capital letter delta
1431	039B	Λ ; Λ	Canadian syllabics pi -> Greek capital letter lamda
14A5	0393	Γ ; Γ	Canadian syllabics mi -> Greek capital letter gamma
14C2	03C3	σ ; σ	Canadian syllabics ni -> Greek small letter sigma
1577	03B4	δ ; δ	Canadian syllabics nunavik ho -> Greek small letter delta

**In Mongolian script confusable**

Input	Output	Display	Comment
1810 *	004F	᠐ ; O	Mongolian digit zero -> Latin capital letter o

**In Han script confusable**

Input	Output	Display	Comment
3007 *	004F	〇 ; O	Ideographic number zero -> Latin capital letter o

Upper case confusable characters are shown for reference, they may not be needed.

**In Latin script confusable, upper case**

Input	Output	Display	Comment
00D0 *	0044 0335	Ð ; Ð	Latin capital eth -> Latin capital d + Combining short stroke overlay
00D8 *	004F 0338	Ø ; Ø	Latin capital letter o with stroke -> Latin capital letter o + Combining long solidus overlay
0110 *	00D0	Ɖ ; Ɖ	Latin capital letter d with stroke -> Latin capital letter eth
0126 *	0048 0335	Ĥ ; Ĥ	Latin capital letter h with stroke -> Latin capital letter h + Combining short stroke overlay
0141 *	004C 0337	Ł ; Ł	Latin capital letter l with stroke -> Latin capital letter l + Combining short solidus overlay
0166	0054 0335	Ƨ ; Ƨ	Latin capital letter t with stroke -> Latin capital letter t + Combining short stroke overlay
0181	0042	Ɓ ; Ɓ	Latin capital letter b with hook -> Latin capital letter b
0182 *	0062 0304	Ɓ̄ ; Ɓ̄	Latin capital letter b with topbar -> Latin small letter b + Combining macron
0187	0043	Ƈ ; Ƈ	Latin capital letter c with hook -> Latin capital letter c
0187 *	0043 02BB	Ƈ ; Ƈ	Latin capital letter c with hook -> Latin capital letter c + modifier letter turned comma
0189 *	00D0	Ɖ ; Ɖ	Latin capital letter African d -> Latin capital letter eth
018A	0044	Ɗ ; Ɗ	Latin capital letter d with hook -> Latin capital letter d
018B *	0064 0304	Ɗ̄ ; Ɗ̄	Latin capital letter d with topbar -> Latin small letter d + Combining macron
0191 *	0046 0321	Ɖ ; Ɖ	Latin capital letter with hook -> Latin capital letter f + Combining palatalized hook below
0193 *	0047 02BB	Ɖ ; Ɖ	Latin capital letter g with hook -> Latin capital letter g + Modifier letter turned comma
0197 *	0049 0335	Ƨ ; Ƨ	Latin capital letter I with stroke -> Latin capital letter I + combining short stroke overlay
0198	004B	Ƒ ; Ƒ	Latin capital letter k with hook -> Latin capital letter k
0198 *	004B 02BB	Ƒ ; Ƒ	Latin capital letter k with hook -> Latin capital letter k + Modifier letter turned comma

019D	004E	Ɔ ; Ɔ	Latin capital letter n with left hook -> Latin capital letter n
019D *	004E 0321	Ɔ ; Ɔ	Latin capital letter n with left hook -> Latin capital letter n + Combining palatalized hook below
019F *	004F 0335	Ɔ ; Ɔ	Latin capital letter o with middle tilde -> Latin capital o + Combining short stroke overlay
01A4	0050	Ɔ ; Ɔ	Latin capital letter p with hook -> Latin capital letter p
01A6	0052	Ɔ ; Ɔ	Latin letter yr -> Latin capital letter r
01AC	0054	Ɔ ; Ɔ	Latin capital letter t with hook -> Latin capital letter t
01AE	0054	Ɔ ; Ɔ	Latin capital letter with retroflex hook -> Latin capital letter t
01AE *	0054 0322	Ɔ ; Ɔ	Latin capital letter with retroflex hook -> Latin capital letter t + Combining retroflex hook below
01B2	0056	Ɔ ; Ɔ	Latin capital letter v with hook -> Latin capital letter v
01B3	0059	Ɔ ; Ɔ	Latin capital letter y with hook -> Latin capital letter y
01B5	005A	Ɔ ; Ɔ	Latin capital letter z with stroke -> Latin capital letter z
01B5 *	005A 0335	Ɔ ; Ɔ	Latin capital letter z with stroke -> Latin capital letter z + Combining short stroke overlay
01BC	0035	5 ; 5	Latin capital letter tone five -> Digit five
01E4 *	0047 0335	Ɔ ; Ɔ	Latin capital letter g with stroke -> Latin capital letter G + Combining short stroke overlay
0224	005A	Ɔ ; Ɔ	Latin capital letter z with hook -> Latin capital letter z
0224 *	005A 0321	Ɔ ; Ɔ	Latin capital letter z with hook -> Latin capital letter z + Combining palatalized hook below

#### In Greek script confusable, upper case

Input	Output	Display	Comment
2206	0394	Δ ; Δ	Increment -> Greek capital letter delta

#### In Cyrillic script confusable, upper case

Input	Output	Display	Comment
0417 *	0033	3 ; 3	Cyrillic capital letter ze -> Digit three
042B *	042C 0049	Ы ; Ы	Cyrillic capital letter yeru -> Cyrillic capital letter soft sign + Latin capital letter i
0472	04E8	Ɔ ; Ɔ	Cyrillic capital letter fita -> Cyrillic capital letter barred o (also confusable with 019F Latin capital letter o with middle tilde)
047C *	0460 0483	Ω ; Ω	Cyrillic capital letter omega with titlo -> Cyrillic capital letter omega + Combining Cyrillic titlo



048C *	0462	Ѣ ; Ѣ	Cyrillic capital letter semisoft sign -> Cyrillic capital letter yat
0490	0413	Ґ ; Ґ	Cyrillic capital letter ghe with upturn -> Cyrillic capital letter ghe (0413 also confusable with 0393 Greek capital letter gamma)
0490 *	0393 02C8	Ґ ; Ґ	Cyrillic capital letter ghe with upturn-> Greek capital letter gamma + Modifier letter vertical line
0492 *	0393 0335	Ғ ; Ғ	Cyrillic capital letter ghe with stroke-> Greek capital letter gamma + Combining short stroke overlay
0496 *	0416 0329	Ҕ ; Ҕ	Cyrillic capital letter zhe with descender -> Cyrillic capital letter zhe + Combining vertical line below
0498 *	0033 0327	Җ ; Җ	Cyrillic capital letter ze -> Digit three + Combining cedilla (TR36 says 0321 Combining palatalized hook below)
049A *	004B 0329	Ҙ ; Ҙ	Cyrillic capital letter ka with descender -> Latin capital letter k + Combining vertical line below
049E *	004B 0335	Ҝ ; Ҝ	Cyrillic capital letter ka with stroke -> Latin capital letter k + Combining short stroke overlay
04A2 *	0048 0329	Ҟ ; Ҟ	Cyrillic capital letter en with tail -> Latin capital letter h + Combining vertical line below
04AA *	0043 0327	Ҝ ; Ҝ	Cyrillic capital letter es with descender -> Latin capital letter c + Combining cedilla (TR 36 says 0322)
04AC *	0054 0329	Ҟ ; Ҟ	Cyrillic capital letter te with descender -> Latin capital letter t + Combining vertical line below
04AE *	0423	Ҟ ; Ҟ	Cyrillic capital letter straight u -> Cyrillic capital letter u) (also confusable with 0059 Latin capital letter y)
04B0 *	0059 0335	Ҟ ; Ҟ	Cyrillic capital letter straight u with stroke -> Latin capital letter y + Combining short stroke overlay (also confusable with 00A5 Yen sign)
04B2 *	0058 0329	Ҟ ; Ҟ	Cyrillic capital letter ha with descender -> Latin capital letter h + Combining vertical line below
04BE *	04BC 0322	Ҟ ; Ҟ	Cyrillic capital letter abkhasian che with descender -> Cyrillic capital letter abkhasian che + Combining retroflex hook below
04C5 *	039B 0321	Ҟ ; Ҟ	Cyrillic capital letter el with tail -> Greek capital letter lamda + Combining palatalized hook below
04C7 *	0048 0321	Ҟ ; Ҟ	Cyrillic capital letter en with hook -> Latin capital letter h + Combining palatalized hook below
04C9 *	0048 0321	Ҟ ; Ҟ	Cyrillic capital letter en with tail -> Latin capital letter h + Combining palatalized hook below
04CB *	04B6	Ҟ ; Ҟ	Cyrillic capital letter khakassian che -> Cyrillic capital letter che with descender
04CD *	004D 0321	Ҟ ; Ҟ	Cyrillic capital letter em with tail -> Latin capital letter m + Combining palatalized hook below
04E8 *	004F 0336	Ҟ ; Ҟ	Cyrillic capital letter barred o -> Latin capital letter o + Combining long stroke overlay

**In mixed script (Greek-Latin) confusable, upper case**

Input	Output	Display	Comment
0391 *	0041	Α ; Α	Greek capital letter alpha -> Latin capital letter a
0392 *	0042	Β ; Β	Greek capital letter beta -> Latin capital letter b
0395 *	0045	Ε ; Ε	Greek capital letter epsilon -> Latin capital letter e

0396 *	005A	Ζ ; Ζ	Greek capital letter zeta -> Latin capital letter z
0397 *	0048	Η ; Η	Greek capital letter eta -> Latin capital letter h
0398	019F	Θ ; Θ	Greek capital letter theta -> Latin capital letter o with middle tilde
0399 *	0049	Ι ; Ι	Greek capital letter iota -> Latin capital letter i
039A *	004B	Κ ; Κ	Greek capital letter kappa -> Latin capital letter k
039C *	004D	Μ ; Μ	Greek capital letter mu -> Latin capital letter m
039D *	004E	Ν ; Ν	Greek capital letter nu -> Latin capital letter n
039F *	004F	Ο ; Ο	Greek capital letter omicron -> Latin capital letter o
03A1 *	0050	Ρ ; Ρ	Greek capital letter rho -> Latin capital letter p
03A3 *	01A9	Σ ; Σ	Greek capital letter sigma -> Latin capital letter esh
03A4 *	0054	Τ ; Τ	Greek capital letter tau -> Latin capital letter t
03A5 *	0059	Υ ; Υ	Greek capital letter upsilon -> Latin capital letter y
03A7 *	0058	Χ ; Χ	Greek capital letter chi -> Latin capital letter x
03DC	0046	Ϝ ; Ϝ	Greek letter digamma -> Latin capital letter f
03E8 *	01A7	Ⲉ ; Ⲉ	Coptic capital letter hori -> Latin capital letter tone two
03F9	0043	Ϛ ; Ϛ	Greek capital lunate sigma symbol -> Latin capital letter c (should be IDN-mapped in a new IDN to capital sigma?)
03FA	004D	ϛ ; ϛ	Greek capital letter san -> Latin capital letter m

**In mixed script (Cyrillic-Latin) confusable, upper case**

Input	Output	Display	Comment
0405 *	0053	Ѕ ; S	Cyrillic capital letter dze -> Latin capital letter s
0406 *	0049	І ; I	Cyrillic capital letter Byelorussian-Ukrainian i -> Latin capital letter i
0408 *	004A	Ј ; J	Cyrillic capital letter je -> Latin capital letter j
0410 *	0041	А ; A	Cyrillic capital letter a -> Latin capital letter a
0411	0062 0304	Б̄ ; B̄	Cyrillic capital letter be -> Latin small letter b + Combining macron
0411 *	0062 0305	Б̅ ; B̅	Cyrillic capital letter be -> Latin small letter b + Combining overline
0412 *	0042	В ; B	Cyrillic capital letter ve -> Latin capital letter b

0415 *	0045	Е ; E	Cyrillic capital letter ie -> Latin capital letter e
041A *	004B	К ; K	Cyrillic capital letter ka -> Latin capital letter k
041C *	004D	М ; M	Cyrillic capital letter m -> Latin capital letter m
041D *	0048	Н ; H	Cyrillic capital letter en -> Latin capital letter h
041E *	004F	О ; O	Cyrillic capital letter o -> Latin capital letter o
0420 *	0050	Р ; P	Cyrillic capital letter er -> Latin capital letter p
0421 *	0043	С ; C	Cyrillic capital letter es -> Latin capital letter c
0422 *	0054	Т ; T	Cyrillic capital letter te -> Latin capital letter t
0423	0059	У ; Y	Cyrillic capital letter u -> Latin capital letter y
0425 *	0058	Х ; X	Cyrillic capital letter ha -> Latin capital letter x
042C	0062	ъ ; b	Cyrillic capital letter soft sign -> Latin small letter b
0474 *	0056	Ѡ ; V	Cyrillic capital letter izhitsa -> Latin capital letter v
048C	0180	Ѣ ; ѣ	Cyrillic capital letter semisoft sign -> Latin small letter b with stroke
0492	0046	Ѥ ; F	Cyrillic capital letter ghe with stroke -> Latin capital letter f
04BA	0068	Ѧ ; h	Cyrillic capital letter shha -> Latin small letter h
04D4 *	00C6	Ѧ ; Æ	Cyrillic capital ligature ae -> Latin capital letter ae
04D8 *	018F	Ѧ ; ə	Cyrillic capital letter schwa -> Latin capital letter schwa
04E0	01B7	Ѧ ; Ʒ	Cyrillic capital letter abkhasian dze
04E8	019F	Ѧ ; Ø	Cyrillic capital letter barred o -> Latin capital letter o with middle tilde
0500	0064	Ѧ ; d	Cyrillic capital letter komi de -> Latin small letter d
050C	0047	Ѧ ; G	Cyrillic capital letter komi sje -> Latin small letter g

**In mixed script (Cyrillic-Greek) confusable, upper case**

Input	Output	Display	Comment
0413 *	0393	Г ; Γ	Cyrillic capital letter barred ghe -> Greek capital letter gamma
041B *	039B	Л ; Λ	Cyrillic capital letter el -> Greek capital letter lamda
041F *	03A0	П ; Π	Cyrillic capital letter pe -> Greek capital letter pi

0424 *	03A6	Ϙ ; Ϙ	Cyrillic capital letter ef -> Greek capital letter phi
--------	------	-------	--

**In mixed script (Armenian-Latin) confusable, upper case**

Input	Output	Display	Comment
0555 *	004F	Օ ; O	Armenian capital letter oh -> Latin capital letter o