Subject: UTS #39 Type values from MSR-2
From: Mark Davis
Date: 2015-05-02
Link: http://goo.gl/fazj7A

I had an action to "Compare the xidmodifications.txt characters in UTS #39 against the ICANN list(s) of maximal starting repertoire characters, and add as appropriate for 8.0." See also icann.org/news/announcement-2-2015-04-27-en.

I added values from MSR-2 that looked pretty straightforward, mapping their reasons for exclusion onto the Type values used in UTS #39 (http://www.unicode.org/reports/tr39/proposed.html#Identifier_Status_and_Type). In my sources, these are marked as "from MSR-2". In general their principles seem to line up with UTS #39's, although there are some differences, because MSR has a narrower scope.

I didn't add any exclusions that didn't have an explicit reason, and didn't add Han or Bopomofo exclusions, since they appear to need more thought. (The reasons for their exclusion from MSR-2 are given in the overview document (icann.org/en/system/files/files/msr-2-overview-14apr15-en.pdf). I also excluded items that were in CLDR exemplar sets, until their status could be investigated. Those are commented out with # below, such as #04CA.

1. I should take an action to prepare a document for 9.0 reviewing the ones I didn't add, and also reviewing any changes in MSR versions over the coming months.
2. *The UTC may want to revert some of these changes, or may wish to defer all of them until 9.0 (allowing more time for review). So I'm listing them here separately for review.*

```
048B        ; Uncommon_Use   # from MSR-2 ( й ) [Ll, Cyrillic] CYRILLIC SMALL LETTER SHORT I WITH TAIL
048D        ; Uncommon_Use   # from MSR-2 ( ҍ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER SEMISOFT SIGN
048F        ; Uncommon_Use   # from MSR-2 ( ҏ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER ER WITH TICK
04C4        ; Uncommon_Use   # from MSR-2 ( ӄ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KA WITH HOOK
04C6        ; Uncommon_Use   # from MSR-2 ( ӆ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EL WITH TAIL
04C8        ; Uncommon_Use   # from MSR-2 ( ӈ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EN WITH HOOK
#04CA       ; Uncommon_Use   # from MSR-2 ( ӊ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EN WITH TAIL
04CE        ; Uncommon_Use   # from MSR-2 ( ӎ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EM WITH TAIL
04F7        ; Uncommon_Use   # from MSR-2 ( ӷ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER GHE WITH DESCENDER
04FB        ; Uncommon_Use   # from MSR-2 ( ӻ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER GHE WITH STROKE AND HOOK
04FD        ; Uncommon_Use   # from MSR-2 ( ӽ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER HA WITH HOOK
04FF        ; Uncommon_Use   # from MSR-2 ( ӿ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER HA WITH STROKE
0511        ; Uncommon_Use   # from MSR-2 ( ԑ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER REVERSED ZE
0513        ; Uncommon_Use   # from MSR-2 ( ԓ ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EL WITH HOOK

08B2        ; Uncommon_Use   # from MSR-2 ( ? ) [Lo, Arabic] ARABIC LETTER ZAIN WITH INVERTED V ABOVE

0A01        ; Uncommon_Use   # from MSR-2 ( ˚  ) [Mn, Gurmukhi] GURMUKHI SIGN ADAK BINDI

#0C01       ; Uncommon_Use   # from MSR-2 ( ఁ ) [Mc, Telugu] TELUGU SIGN CANDRABINDU

0F6A..0F6C  ; Uncommon_Use   # from MSR-2 (   ..   ) [Lo.., Tibetan..] TIBETAN LETTER FIXED-FORM RA..
0F82..0F83  ; Uncommon_Use   # from MSR-2 ( ˚..˚ ) [Mn.., Tibetan..] TIBETAN SIGN NYI ZLA NAA DA..
0FAE..0FB0  ; Uncommon_Use   # from MSR-2 (   ..   ) [Mn.., Tibetan..] TIBETAN SUBJOINED LETTER ZHA..
0FC6        ; Uncommon_Use   # from MSR-2 (   ) [Mn, Tibetan] TIBETAN SYMBOL PADMA GDAN

1028        ; Uncommon_Use   # from MSR-2 (   ) [Lo, Myanmar] MYANMAR LETTER MON E
1033..1034  ; Uncommon_Use   # from MSR-2 (   ..   ) [Mn.., Myanmar..] MYANMAR VOWEL SIGN MON II..
1050..1060  ; Uncommon_Use   # from MSR-2 (   ..   ) [Lo.., Myanmar..] MYANMAR LETTER SHA..
1065..1074  ; Uncommon_Use   # from MSR-2 (   ..   ) [Lo.., Myanmar..] MYANMAR LETTER WESTERN PWO KAREN THA..
108E..108F  ; Uncommon_Use   # from MSR-2 (   ..   ) [Lo.., Myanmar..] MYANMAR LETTER RUMAI PALAUNG FA..
109A..109D  ; Uncommon_Use   # from MSR-2 (   ..   ) [Mc.., Myanmar..] MYANMAR SIGN KHAMTI TONE-1..
```

```
10F9..10FA ; Uncommon_Use    # from MSR-2 ( ჹ..ჺ ) [Lo.., Georgian..] GEORGIAN LETTER TURNED GAN..
10FD..10FF ; Uncommon_Use    # from MSR-2 (  ..  ) [Lo.., Georgian..] GEORGIAN LETTER AEN..

2DA0..2DA6 ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE SSA..
2DA8..2DAE ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE CCA..
2DB0..2DB6 ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE ZZA..
2DB8..2DBE ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE CCHA..
2DC0..2DC6 ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE QYA..
2DC8..2DCE ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE KYA..
2DD0..2DD6 ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE XYA..
2DD8..2DDE ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE GYA..
AB11..AB16 ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE DZU..
AB20..AB26 ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE CCHHA..
AB28..AB2E ; Uncommon_Use    # from MSR-2 (   ..  ) [Lo.., Ethiopic..] ETHIOPIC SYLLABLE BBA..

0201       ; Technical       # from MSR-2 ( ȁ ) [Ll, Latin] LATIN SMALL LETTER A WITH DOUBLE GRAVE
0203       ; Technical       # from MSR-2 ( ȃ ) [Ll, Latin] LATIN SMALL LETTER A WITH INVERTED BREVE
0205       ; Technical       # from MSR-2 ( ȅ ) [Ll, Latin] LATIN SMALL LETTER E WITH DOUBLE GRAVE
0207       ; Technical       # from MSR-2 ( ȇ ) [Ll, Latin] LATIN SMALL LETTER E WITH INVERTED BREVE
0209       ; Technical       # from MSR-2 ( ȉ ) [Ll, Latin] LATIN SMALL LETTER I WITH DOUBLE GRAVE
020B       ; Technical       # from MSR-2 ( ȋ ) [Ll, Latin] LATIN SMALL LETTER I WITH INVERTED BREVE
020D       ; Technical       # from MSR-2 ( ȍ ) [Ll, Latin] LATIN SMALL LETTER O WITH DOUBLE GRAVE
020F       ; Technical       # from MSR-2 ( ȏ ) [Ll, Latin] LATIN SMALL LETTER O WITH INVERTED BREVE
0211       ; Technical       # from MSR-2 ( ȑ ) [Ll, Latin] LATIN SMALL LETTER R WITH DOUBLE GRAVE
0213       ; Technical       # from MSR-2 ( ȓ ) [Ll, Latin] LATIN SMALL LETTER R WITH INVERTED BREVE
0215       ; Technical       # from MSR-2 ( ȕ ) [Ll, Latin] LATIN SMALL LETTER U WITH DOUBLE GRAVE
0217       ; Technical       # from MSR-2 ( ȗ ) [Ll, Latin] LATIN SMALL LETTER U WITH INVERTED BREVE
1E01       ; Technical       # from MSR-2 ( ḁ ) [Ll, Latin] LATIN SMALL LETTER A WITH RING BELOW
1E13       ; Technical       # from MSR-2 ( ḓ ) [Ll, Latin] LATIN SMALL LETTER D WITH CIRCUMFLEX BELOW
1E19       ; Technical       # from MSR-2 ( ḙ ) [Ll, Latin] LATIN SMALL LETTER E WITH CIRCUMFLEX BELOW
1E1B       ; Technical       # from MSR-2 ( ḛ ) [Ll, Latin] LATIN SMALL LETTER E WITH TILDE BELOW
1E2B       ; Technical       # from MSR-2 ( ḫ ) [Ll, Latin] LATIN SMALL LETTER H WITH BREVE BELOW
1E2D       ; Technical       # from MSR-2 ( ḭ ) [Ll, Latin] LATIN SMALL LETTER I WITH TILDE BELOW
1E3D       ; Technical       # from MSR-2 ( ḽ ) [Ll, Latin] LATIN SMALL LETTER L WITH CIRCUMFLEX BELOW
1E4B       ; Technical       # from MSR-2 ( ṋ ) [Ll, Latin] LATIN SMALL LETTER N WITH CIRCUMFLEX BELOW
1E71       ; Technical       # from MSR-2 ( ṱ ) [Ll, Latin] LATIN SMALL LETTER T WITH CIRCUMFLEX BELOW
1E73       ; Technical       # from MSR-2 ( ṳ ) [Ll, Latin] LATIN SMALL LETTER U WITH DIAERESIS BELOW
1E75       ; Technical       # from MSR-2 ( ṵ ) [Ll, Latin] LATIN SMALL LETTER U WITH TILDE BELOW
1E77       ; Technical       # from MSR-2 ( ṷ ) [Ll, Latin] LATIN SMALL LETTER U WITH CIRCUMFLEX BELOW

030F..0311 ; Technical       # from MSR-2 ( ̏ .. ̑ ) [Mn.., Inherited..] COMBINING DOUBLE GRAVE ACCENT..
0313..0314 ; Technical       # from MSR-2 ( ̓ .. ̔ ) [Mn.., Inherited..] COMBINING COMMA ABOVE..
0324..0325 ; Technical       # from MSR-2 ( ̤ .. ̥ ) [Mn.., Inherited..] COMBINING DIAERESIS BELOW..
032D..032E ; Technical       # from MSR-2 ( ̭ .. ̮ ) [Mn.., Inherited..] COMBINING CIRCUMFLEX ACCENT BELOW..
0330       ; Technical       # from MSR-2 ( ̰ ) [Mn, Inherited] COMBINING TILDE BELOW
0335       ; Technical       # from MSR-2 ( ̵ ) [Mn, Inherited] COMBINING SHORT STROKE OVERLAY
0338..0339 ; Technical       # from MSR-2 ( ̸ .. ̹ ) [Mn.., Inherited..] COMBINING LONG SOLIDUS OVERLAY..

0342       ; Technical       # from MSR-2 ( ͂ ) [Mn, Greek] COMBINING GREEK PERISPOMENI
03FC       ; Technical       # from MSR-2 ( ϼ ) [Ll, Greek] GREEK RHO WITH STROKE SYMBOL

0529       ; Technical       # from MSR-2 ( ? ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EN WITH LEFT HOOK
052F       ; Technical       # from MSR-2 ( ? ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EL WITH DESCENDER

0559       ; Technical       # from MSR-2 ( ՙ ) [Lm, Armenian] ARMENIAN MODIFIER LETTER LEFT HALF RING

06AC       ; Technical       # from MSR-2 ( ڬ ) [Lo, Arabic] ARABIC LETTER KAF WITH DOT ABOVE
06E5..06E6 ; Technical       # from MSR-2 ( ۥ .. ۦ ) [Lm.., Arabic..] ARABIC SMALL WAW..
08A1       ; Technical       # from MSR-2 ( ? ) [Lo, Arabic] ARABIC LETTER BEH WITH HAMZA ABOVE

0653       ; Technical       # from MSR-2 (  ) [Mn, Arabic|Syriac] ARABIC MADDAH ABOVE
```

#0950   ; Technical   # from MSR-2 (ॐ) [Lo, Devanagari] DEVANAGARI OM

097D   ; Technical   # from MSR-2 (?) [Lo, Devanagari] DEVANAGARI LETTER GLOTTAL STOP

#0A74   ; Technical   # from MSR-2 (ੴ) [Lo, Gurmukhi] GURMUKHI EK ONKAR

#0AD0   ; Technical   # from MSR-2 (ૐ) [Lo, Gujarati] GUJARATI OM

0B82   ; Technical   # from MSR-2 (ஂ) [Mn, Tamil] TAMIL SIGN ANUSVARA
0BD0   ; Technical   # from MSR-2 (ௐ) [Lo, Tamil] TAMIL OM

#0E2F   ; Technical   # from MSR-2 (๏) [Lo, Thai] THAI CHARACTER PAIYANNOI

#0EAF   ; Technical   # from MSR-2 (ຯ) [Lo, Lao] LAO ELLIPSIS

0F00   ; Technical   # from MSR-2 (ༀ) [Lo, Tibetan] TIBETAN SYLLABLE OM
0F35   ; Technical   # from MSR-2 (༵) [Mn, Tibetan] TIBETAN MARK NGAS BZUNG NYI ZLA
0F37   ; Technical   # from MSR-2 (༷) [Mn, Tibetan] TIBETAN MARK NGAS BZUNG SGOR RTAGS
0F3E..0F3F ; Technical   # from MSR-2 (༾..༿) [Mc.., Tibetan..] TIBETAN SIGN YAR TSHES..
0F7B   ; Technical   # from MSR-2 (ཻ) [Mn, Tibetan] TIBETAN VOWEL SIGN EE
0F7D   ; Technical   # from MSR-2 (ཽ) [Mn, Tibetan] TIBETAN VOWEL SIGN OO

3007   ; Technical   # from MSR-2 (〇) [Nl, Han] IDEOGRAPHIC NUMBER ZERO

3099..309A ; Technical   # from MSR-2 (゙.. ゚) [Mn.., Hiragana|Katakana..] COMBINING KATAKANA-HIRAGANA
VOICED SOUND MARK..

0138   ; Obsolete   # from MSR-2 (ĸ) [Ll, Latin] LATIN SMALL LETTER KRA

037B..037D ; Obsolete   # from MSR-2 (ͻ..ͽ) [Ll.., Greek..] GREEK SMALL REVERSED LUNATE SIGMA SYMBOL..

049D   ; Obsolete   # from MSR-2 (ҝ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KA WITH VERTICAL STROKE
04A7   ; Obsolete   # from MSR-2 (ҧ) [Ll, Cyrillic] CYRILLIC SMALL LETTER PE WITH MIDDLE HOOK
04B9   ; Obsolete   # from MSR-2 (ҹ) [Ll, Cyrillic] CYRILLIC SMALL LETTER CHE WITH VERTICAL STROKE
0501   ; Obsolete   # from MSR-2 (ԁ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI DE
0503   ; Obsolete   # from MSR-2 (ԃ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI DJE
0505   ; Obsolete   # from MSR-2 (ԅ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI ZJE
0507   ; Obsolete   # from MSR-2 (ԇ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI DZJE
0509   ; Obsolete   # from MSR-2 (ԉ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI LJE
050B   ; Obsolete   # from MSR-2 (ԋ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI NJE
050D   ; Obsolete   # from MSR-2 (ԍ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI SJE
050F   ; Obsolete   # from MSR-2 (ԏ) [Ll, Cyrillic] CYRILLIC SMALL LETTER KOMI TJE
0515   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER LHA
0517   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER RHA
0519   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER YAE
051B   ; Obsolete   # from MSR-2 (ԛ) [Ll, Cyrillic] CYRILLIC SMALL LETTER QA
051D   ; Obsolete   # from MSR-2 (ԝ) [Ll, Cyrillic] CYRILLIC SMALL LETTER WE
051F   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER ALEUT KA
0521   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EL WITH MIDDLE HOOK
0523   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER EN WITH MIDDLE HOOK
0527   ; Obsolete   # from MSR-2 (   ) [Ll, Cyrillic] CYRILLIC SMALL LETTER SHHA WITH DESCENDER

063B..063F ; Obsolete   # from MSR-2 (..   ) [Lo.., Arabic..] ARABIC LETTER KEHEH WITH TWO DOTS ABOVE..
077E..077F ; Obsolete   # from MSR-2 (..   ) [Lo.., Arabic..] ARABIC LETTER SEEN WITH INVERTED V..

#093D   ; Obsolete   # from MSR-2 (ऽ) [Lo, Devanagari] DEVANAGARI SIGN AVAGRAHA

0960..0963 ; Obsolete   # from MSR-2 (ॠ..ॣ) [Lo.., Devanagari..] DEVANAGARI LETTER VOCALIC RR..

0971   ; Obsolete   # from MSR-2 (˙) [Lm, Devanagari] DEVANAGARI SIGN HIGH SPACING DOT

#09BD   ; Obsolete   # from MSR-2 (২) [Lo, Bengali] BENGALI SIGN AVAGRAHA
#09E0..09E3 ; Obsolete   # from MSR-2 (ঋ.ৣ   ) Lo.., Bengali..] BENGALI LETTER VOCALIC RR..

#0ABD   ; Obsolete   # from MSR-2 (ऽ) [Lo, Gujarati] GUJARATI SIGN AVAGRAHA

#0AE0 ; Obsolete   # from MSR-2 (ૠૢ   ]Lo.., Gujarati..] GUJARATI LETTER VOCALIC RR..

0AE1..0AE3 ; Obsolete   # from MSR-2 (ૠૢ   ]Lo.., Gujarati..] GUJARATI LETTER VOCALIC RR..

0B3D   ; Obsolete   # from MSR-2 (ଽ) [Lo, Oriya] ORIYA SIGN AVAGRAHA
0B60..0B61 ; Obsolete   # from MSR-2 (ଡ଼..ଡ଼) [Lo.., Oriya..] ORIYA LETTER VOCALIC RR..

0C3D   ; Obsolete   # from MSR-2 (ఽ) [Lo, Telugu] TELUGU SIGN AVAGRAHA
#0C60..0C61 ; Obsolete   # from MSR-2 (ఋ..ఌ) [Lo.., Telugu..] TELUGU LETTER VOCALIC RR..

#0CBD   ; Obsolete   # from MSR-2 (ಽ) [Lo, Kannada] KANNADA SIGN AVAGRAHA
#0CE0..0CE3 ; Obsolete   # from MSR-2 (ಖೄ   Lo.., Kannada..] KANNADA LETTER VOCALIC RR..
0CE2..0CE3 ; Obsolete   # from MSR-2 (ಖೄ   Lo.., Kannada..] KANNADA LETTER VOCALIC RR..
0CF1..0CF2 ; Obsolete   # from MSR-2 (ೱ..ೲ) [Lo.., Kannada..] KANNADA SIGN JIHVAMULIYA..

0D3A   ; Obsolete   # from MSR-2 (   ) [Lo, Malayalam] MALAYALAM LETTER TTTA
0D3D   ; Obsolete   # from MSR-2 (∫) [Lo, Malayalam] MALAYALAM SIGN AVAGRAHA
#0D4C   ; Obsolete   # from MSR-2 (� ൗ) [Mc, Malayalam] MALAYALAM VOWEL SIGN AU
0D4E   ; Obsolete   # from MSR-2 (   ) [Lo, Malayalam] MALAYALAM LETTER DOT REPH
#0D60..0D61 ; Obsolete   # from MSR-2 (ൠ..ൡ) [Lo.., Malayalam..] MALAYALAM LETTER VOCALIC
RR..

0F86..0F8F ; Obsolete   # from MSR-2 (ཆ̔..   ) [Mn.., Tibetan..] TIBETAN SIGN LCI RTAGS..

17DC   ; Obsolete   # from MSR-2 (   ) [Lo, Khmer] KHMER SIGN AVAKRAHASANYA