

DOC TYPE: Working Group Document
TITLE: Proposed codepoints and properties for characters in “Proposal to encode Arabic characters used for Bashkir, Belarusian, Crimean Tatar, Karachay, Karakalpak, and Tatar languages”
SOURCE: Roozbeh Pournader (via D. Anderson, UC Berkeley)
ACTION: For consideration by WG2
DISTRIBUTION: ISO/IEC JTC1/SC2/WG

This document summarizes the names, codepoints, and properties for 5 Arabic characters originally proposed in “Proposal to encode Arabic characters used for Bashkir, Belarusian, Crimean Tatar, Karachay, Karakalpak, and Tatar languages” (L2/11-138), which were approved by Unicode Technical Committee 13 May 2011.

08AD;ARABIC LETTER LOW ALEF;Lo;0;AL;;;;;N;;;;;
08AE;ARABIC LETTER DAL WITH THREE DOTS BELOW;Lo;0;AL;;;;;N;;;;;
08AF;ARABIC LETTER SAD WITH THREE DOTS BELOW;Lo;0;AL;;;;;N;;;;;
08B0;ARABIC LETTER GAF WITH INVERTED STROKE;Lo;0;AL;;;;;N;;;;;
08B1;ARABIC LETTER STRAIGHT WAW;Lo;0;AL;;;;;N;;;;;

Data for ArabicShaping.txt:

08AD; LOW ALEF; U; No_Joining_Group
08AE; DAL WITH 3 DOTS BELOW; R; DAL
08AF; SAD WITH 3 DOTS BELOW; D; SAD
08B0; KEHEH WITH STROKE BELOW; D; GAF
08B1; STRAIGHT WAW; R; STRAIGHT WAW

Note: The other three characters proposed in L2/11-138 are already representable using the following sequences:

Waw with madda above: <U+0648, U+0653>
Farsi yeh with hamza below: <U+0649, U+0655>
Farsi yeh with damma above: <U+0649, U+064F>

(Note that the latter two use an ALEF MAKSURA as the first character, which is a dual-joining letter.)