2010-10-27

Title: Refixing the bidi class for U+070F SYRIAC ABBREVIATION MARK

Author: Roozbeh Pournader (HighTech Passport) and Behdad Esfahbod (Google)

Action: For consideration by the UTC

Date: 2010-10-27

Introduction

In the November 2009 meeting of the UTC, it was decided to change the bidi class for U+070F SYRIAC ABBREVIATION MARK from BN (Boundary Neutral) to AL (Right-to-Left Arabic). This was consensus 121-C28. The change was requested by Mark Davis (L2/09-313) and the new bidi class was suggested by Roozbeh Pournader.

This bidi class is the best option, since the character is most commonly followed by a string of Syriac letters (also of bidi class AL). Such a class would make sure that the bidi run consisting of this character and the Syriac letters following it are not broken and most rendering engines would have an easier time rendering the combination of the character and the letter sequence following it.

This is the same pattern followed for Arabic subtending marks U+0600..0604 and the enclosing mark U+06DD: these are most commonly followed by digits, characters that are either of the bidi class AN (Arabic Number) or are of bidi class EN (European Number) but will change their class to AN in the middle of the bidirectional algorithm (step W2) because the subtending mark is preceded by a character of the type AL (Right-to-Left Arabic). For more details, see L2/06-332.

Unfortunately, the UTC consensus was incorrectly reflected in the Unicode Character Database. Unicode 6.0 data files mistakenly changed the bidi class of this character to AN (Arabic Number) instead of AL (Right-to-Left Arabic). The authors believe this was just human error considering that the other similar character were of the bidi type AN (Arabic Number).

Request

• Change the bidi class of U+070F SYRIAC ABBREVIATION MARK in the UCD to AL (Right-to-Left Arabic) to reflect the UTC consensus 121-C28.