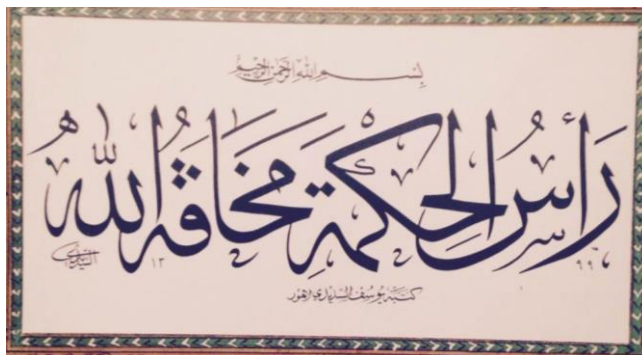# Variant Analysis for Arabic Script Code Point Repertoire
for LGR for the Root Zone
ver. 0.8

General Note: We consider fonts which follow general Arabic calligraphy principles and font sizes which are normally used. One can always make fonts stylistically ambiguous (by not following calligraphy conventions) and font size very small to make text difficult to read. In latter cases, any arbitrary letters can become confusing in any script.

Send feedback to: tf-aidn@meswg.org by 10 January 2015.

Variant Principles

1. Two code points are variants if they are visually same to each other in any one of the four positions/forms (isolated, initial, medial or final form).

2. Two code points are variants if they are visually similar and some Arabic script using communities use the two as stylistic variations of each other, within the calligraphic tradition.

3. Two code points are variants if they are visually similar so that Arabic script using communities may consider the two as stylistic variations of each other.

4. Two code points are variants if they are only different in their dot orientation, i.e. dots are rotated but their count and placement (above or below) does not change. Dot orientation is used in calligraphic tradition to fit is dots in tighter spaces, as shown in the figure below. Note that the ة in مخافة is written with dots rotated vertically (due to constraint in horizontal space).

5. Code points representing same letter in Western (African) Arabic orthography and the Conventional Arabic orthography are considered variants as Arabic script readers from the two orthographic traditions may confuse such letters. Examples are illustrated below.

Conventional Arabic Orthography	**Western (African) Arabic Orthography**

6. Variant sets are symmetric (if A is a variant of B, then vice versa is also true) and transitive (if A is a variant of B and B is a variant of C, then A is also a variant of C).

Variant Sets

Each table represents the set of code points which form a variant set. The tables show the form of each of the characters in all the four forms (initial, medial, final and isolated). All variants are allocatable because each of these letters are used by some language within the Arabic script. Any collocation restrictions within a label may be managed through Whole Label Evaluation rules, if needed. The final column states the applicable principle for which the code point is included in the variant set.

Table 1: YEH Class

Unicode Code Point	Isolated Form	Initial Form	Final Form	Medial Form	Applicable Principle No.
0649	ى	ىـ	ـى	ـىـ	1
06CC	ى	يـ	ـى	ـيـ	1
064A	ي	يـ	ـي	ـيـ	1
06D0	ې	ېـ	ـې	ـېـ	4
067B	ٻ	ٻـ	ـٻ	ـٻـ	4
06CD	ؽ		ـؽ		3
06D2	ے		ـے		5 (0649)

Table 2: HEH Class

06C1	ہ	ہـ	ـہ	ـہـ	1
06D5	ە		ـە		1
0647	ه	هـ	ـه	ـهـ	1
06BE	ھ	ھـ	ـھ	ـھـ	1

Table 3: KAF Class

0643	ك	كب	بك	بكب	1
06A9	ک	کب	بک	بکب	1
06AA	ڪ	ڪب	بڪ	بڪب	2, 5 (0643)

Table 4: PEH Class

067E	پ	پب	بپ	بپب	1
0752	ٻ	ٻب	بٻ	بٻب	4
06D1	ؠ	پب	بؠ	بپب	1
06BD	ڽ	پب	بڽ	بپب	1

Table 5: NOON WITH TWO DOTS BELOW Class

0754	ݔ	ݔب	بݔ	بݔب	1
0767	ݧ	ݧب	بݧ	بݧب	1
08A9	ࢩ	ب⍰	⍰ب	ب⍰ب	1

0641	ف	فب	بف	بفب	1
06A2	ڢ	ڢب	بڢ	بڢب	5 (0641)
06A7	ڧ	ڧب	بڧ	بڧب	1
0642	ق	قب	بق	بقب	5 (06A7)

4

Table 7: TEH MARBUTA Class

0629	ة		بة		1
06C3	ة		بَة		1

Table 8: TTHEH Class

0679	ٹ	ٹب	بٹ	بٹب	1
06BB	ٻ	ٻب	بٻ	بٻب	1

Table 9: NOON Class

0646	ن	نب	بن	بنب	1
06BA	ں	ںب	بں	بںب	1, 5 (0646)

Table 10: TEH Class

062A	ت	تب	بت	بتب	4
067A	ث	ثب	بث	بثب	4

Table 11: NYEH Class

0683	ج	چب	بج	بچب	4
0684	ج	چب	بج	بچب	4

Table 12: VEH Class

06A4	ڤ	ڤب	بڤ	بڤب	1
06A8	ڨ	ڨب	بڨ	بڨب	1

Table 13: THEH Class

062B	ث	ثب	بث	بثب	4
067D	ٿ	ٿب	بٿ	بٿب	4

Table 14: NG Class

06AD	ڭ	ڭب	بڭ	بڭب	1
0763	ڳ	ڳب	بڳ	بڳب	1

Table 15: DUL Class

068E	ڎ		بڎ		4
068F	ڏ		بڏ		4

Table 16: ALEF with HAMZA Class

0623	أ		بأ		3
0672	ٲ		بٲ		3

Additional Question:

Though Hafs and Wars are used for traditions associated with recitation and reading. However, in recent literature the conventional Arabic orthography and western (African) Arabic orthography are also referred to as Hafs and Wars[1]. Is it appropriate to use these names for these orthographic differences as well?

---

[1] See http://scriptsource.org/cms/scripts/page.php?item_id=entry_detail&uid=yq85dn9q24 and

http://www.unicode.org/L2/L2014/14207-warsh-model.pdf for further details on Wars orthography.