Title: Proposal to encode Al-Dani Quranic marks used in Quran published in Libya

Action: For consideration by UTC and ISO/IEC JTC1/SC2/WG2

Author: Lateef Sagar Shaikh

Date: 16-Feb-2016

Introduction

In Proposal L2/16-044 by Abudena, the author has identified all marks which are used in Quran in Libya. These marks were introduced by Uthman bin Said bin Uthman Ibn Omar Abu Amr Al-Dani.

In this proposal I have tried to simplify the initial proposal L2/16-044 so that we can achieve same rendering with less encoding challenge.

Requests

The author requests the encoding of alternate dammatan which is used in Quranic text written and published in Pakistan.

Character name and shape

#	Shape	Codepoint	Name	Reference image
1		U+08BE	ARABIC ALTERNATE DAMMA	1
2		U+08BF	ARABIC ALTERNATE DAMMATAN 1	1
3		U+08C0	ARABIC ALTERNATE DAMMATAN 2	2
4		U+08C1	ARABIC ALTERNATE FATHATAN	3
5		U+08C2	ARABIC ALTERNATE KASRATAN	4
6		U+08C3	ARABIC ALEF MOKHASES	4

7	ٓ◌	U+08C4	ARABIC LETTER ALEF MOKHASES WITH MADDAH ABOVE	5
8	◌ص	U+08C5	ARABIC ALTERNATE SIGN OF WAQAF	1
9	◌ىَ	U+08C6	ARABIC SMALL HIGH YEH WITH FATHA ABOVE	6
10	◌و	U+08C7	ARABIC SMALL WAW BELOW	7
11	◌ٮٟ	U+08C8	ARABIC LETTER HIGH YEH MAQUSA WITH DOT ABOVE	8
12	◌ٜ	U+08C9	ARABIC LETTER SUKUN BELOW	10
13	ـۥـ	U+08CA	ARABIC SMALL WAW INSIDE TATWEEL	9
14	ـۦـ	U+08CB	ARABIC HAMZA INSIDE TATWEEL	9
15	ٱ	U+08CC	ARABIC ALTERNATE ALEF WASLA	1
16	◌ٖ	U+08CD	ARABIC ALEF MOTHAFAR	5

Figures

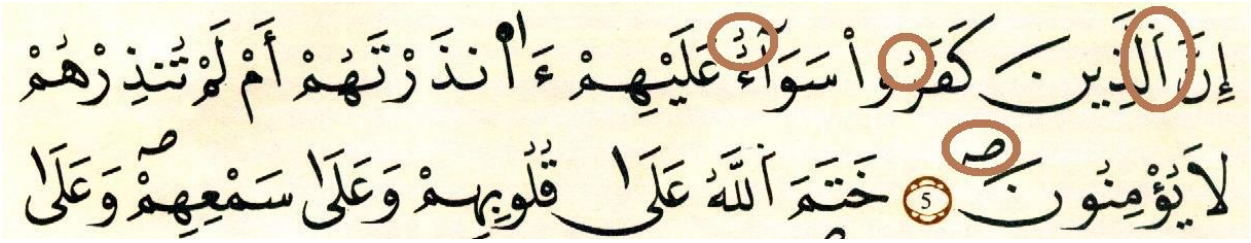


Figure 1

Figure 2

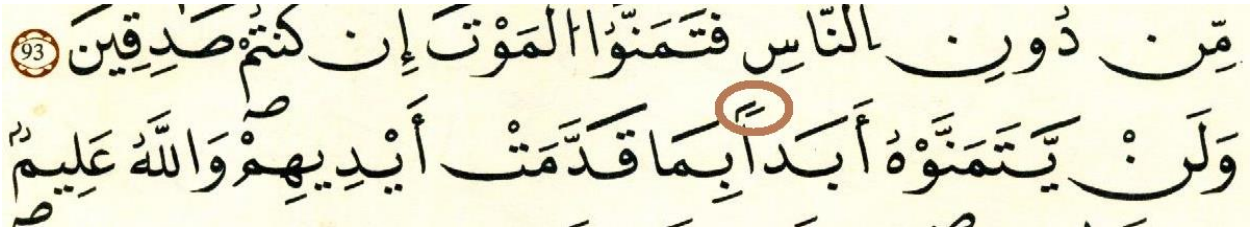


Figure 3

Figure 4

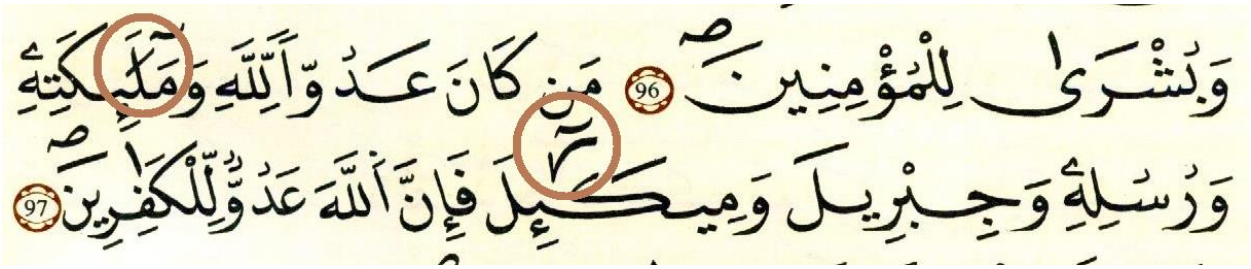


Figure 5

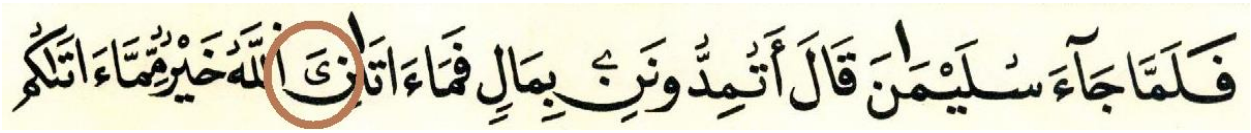


Figure 6

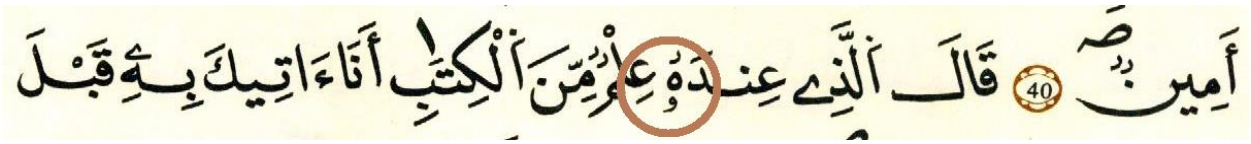


Figure 7

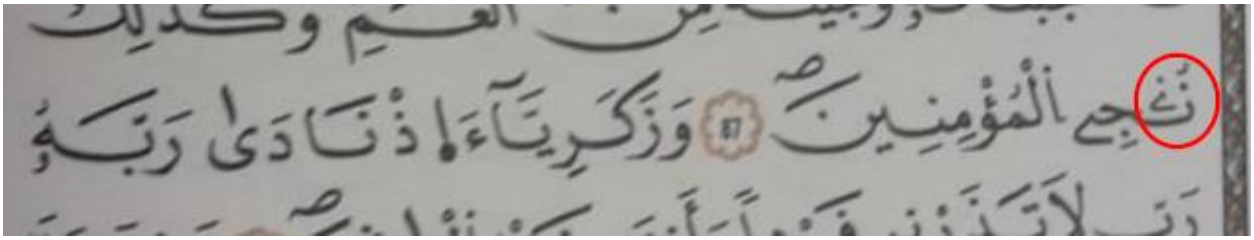


Figure 8

Figure 9

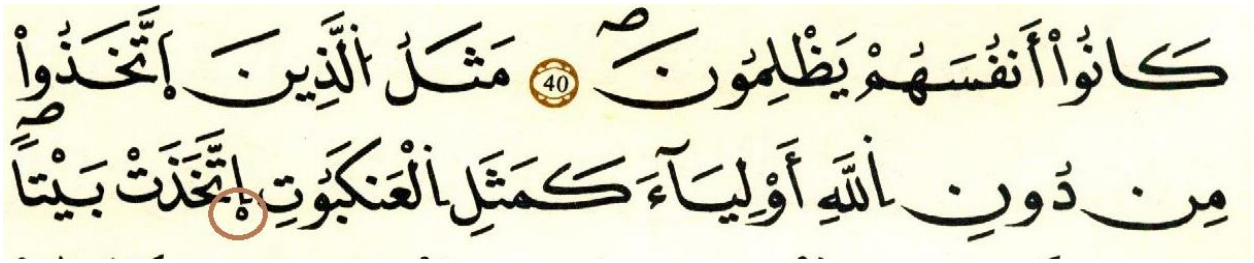


Figure 10

Shapes of suggested characters

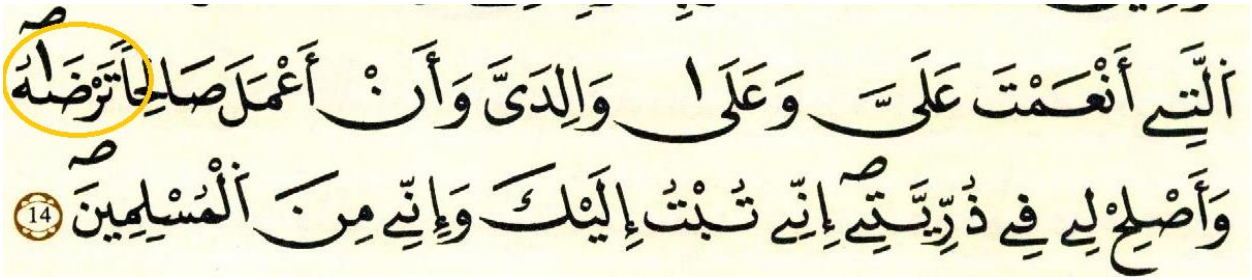
Character	Isolated	Final	Middle	Initial
۞	۞		۞	
ۜ	ۜ		ۜ	

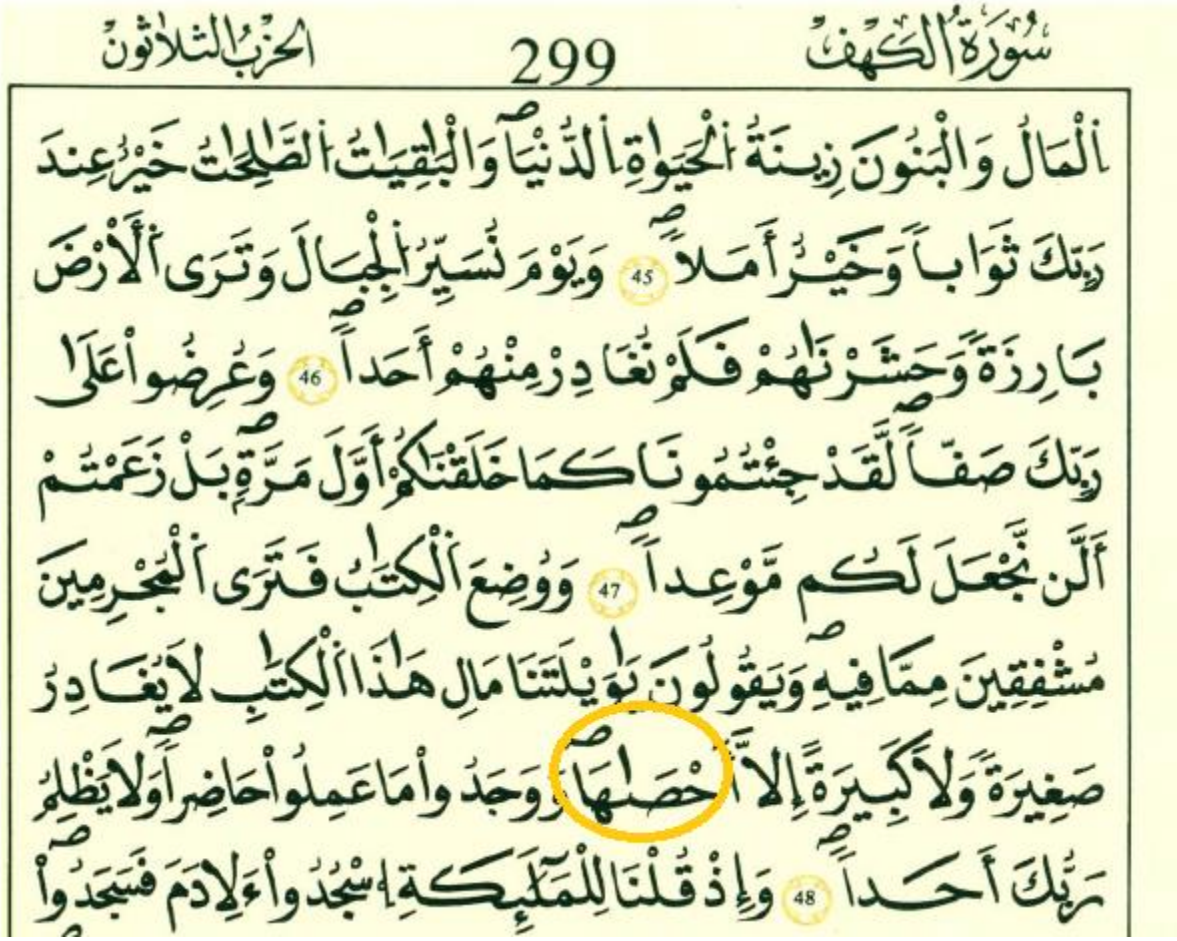
ا	ا	ـا		
ـَ	ـَ	ـَ		

Other symbols in L2/16-044

Sad and Dad without Senna:

There are instances of Sad and Dad in the Holy Quran written Al-Dani style where Alef Khingarih appears on Sad and Dad and they have same shape as regular Arabic Sad and Dad, as shown in the following figures:





Which means that removing senna is a calligraphic style and should be programmed in the font if required.

Alef Khingarih:

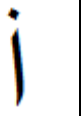
This character has same visual and linguistic properties are “Arabic Letter Superscript Alef – U 670”. Therefore this should not be encoded as a separate character.

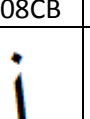
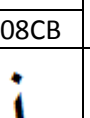
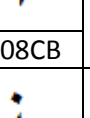
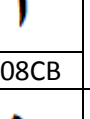
Alef with connected marks:

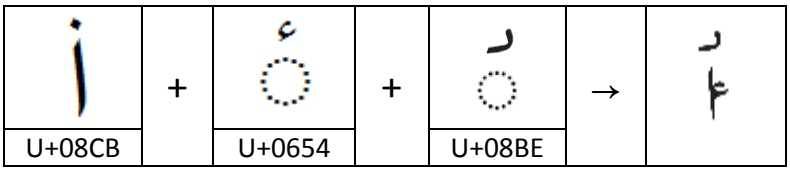
I am proposing a new Alef Wasla to be encoded and it can programmed in the font that whenever this Alef has a fatha, kasra or other symbol then that symbol should be connected. This will make it easy to have Alef with connected symbols and without symbols.

In the following table I have shown how these transformations can be implemented:

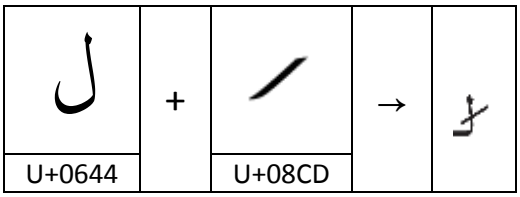
ٱ	+	◌َ	→	أَ
U+08CB		U+064E		

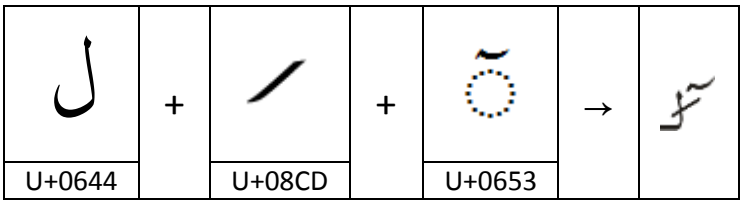
	+		→	
U+08CB		U+0650		
	+		→	
U+08CB		U+064F		

	+		+		→	
U+08CB		U+0652		U+064E		
	+		+		→	
U+08CB		U+0652		U+0650		
	+		+		→	
U+08CB		U+0652		U+064F		
	+		+		→	
U+08CB		U+08C9		U+064E		
	+		+		→	
U+08CB		U+08C9		U+0650		
	+		+		→	
U+08CB		U+08C9		U+064F		



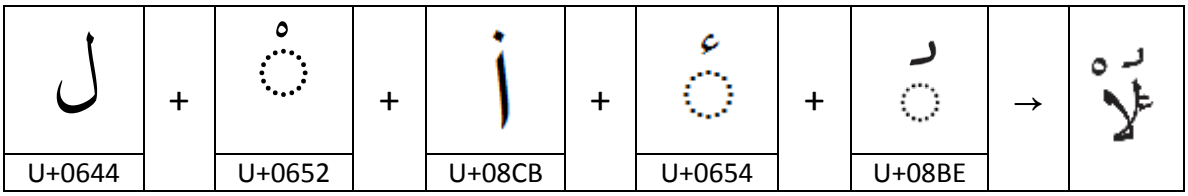
Laam with Alef Mothafar

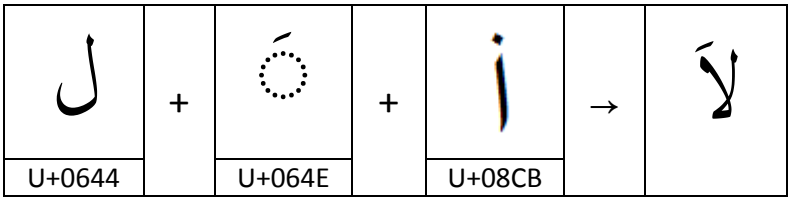


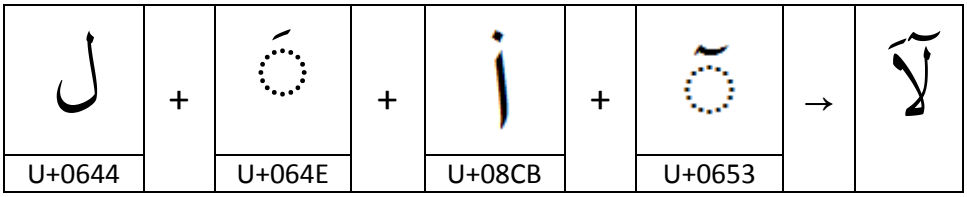


Laam with Alef ligatures

When marks are placed over Laam Alef ligatures nn Al-Dani style, then right side stroke is considered as Alef and left side stroke is considered as Laam. The Alternate Alef Wasla suggested in this proposal can help in solving this issue:







ل	+	ّ	+	َ	+	ا	→	لاَّ
U+0644		U+0651		U+064E		U+08CB		

ل	+	ٗ	+	ا	+	ٔ	+	َ	→	لأَ
U+0644		U+0652		U+08CB		U+0654		U+064E		
ل	+	ٗ	+	ا	+	ٕ	+	ِ	→	لإِ
U+0644		U+0652		U+08CB		U+0655		U+0650		

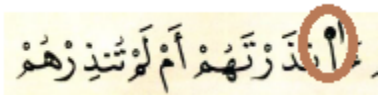
ل	+	َ	+	ـ	+	ٔ	+	َ	+	ل	+	ٗ	+	ا	→	لَ
U+0644		U+064E		U+0640		U+0654		U+064E		U+0644		U+0652		U+08CB		

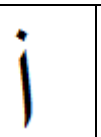
Yeh with small higher yeh

This combination is already supported therefore there is no need to do anything.

ي	+	ِ	+	ۧ	+	ّ	→	يِّ
U+064A		U+0650		U+06E7		U+0651		

Not in original proposal



	+		→	
U+08CB		U+06EC		

Reference

1. Proposal to encode Quranic marks used in Quran published in Libya (Narration of Qaloon with script Aldani)

http://www.unicode.org/L2/L2016/16044-aldani-with-cmts.pdf

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** Proposal to encode Al-Dani Quranic Marks used in Quran published in Libya
2. Requester's name: *Lateef Sagar Shaikh*
3. Requester type (Member body/Liaison/Individual contribution): *Individual Contribution*
4. Submission date: *2/16/2016*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: ✓
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters):
      Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block: ✓
      Name of the existing block: *Arabic*
2. Number of characters in proposal: *16*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary | B.1-Specialized (small collection) ✓ | B.2-Specialized (large collection)
   C-Major extinct | D-Attested extinct | E-Minor extinct
   F-Archaic Hieroglyphic or Ideographic | G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? *Yes*
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *Yes*
   b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
   Lateef Sagar Shaikh
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
   Lateef Sagar Shaikh lateef_sagar@yahoo.com
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *Yes*
7. Special encoding issues:
   Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *Yes*
   See the section titled Character name and shape

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[1] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *No*
   If YES explain
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? *No*
   If YES, with whom?
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *Yes*
   Reference: *Muslims in North Africa*
4. The context of use for the proposed characters (type of use; common or rare) *Quranic*
   Reference: *The proposed mark is very common in Quran published in Libya*
5. Are the proposed characters in current use by the user community? Yes
   If YES, where? Reference: *Used by Muslins in North Africa*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *Yes*
   If YES, is a rationale provided? *Yes*
   If YES, reference: *Character could be kept with their counterparts in Arabic block*
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *Yes*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? *No*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *Yes*
    If YES, is a rationale for such use provided? *Yes*
    If YES, reference: *Similarity to already encoded characters*
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? *No*
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *No*
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility characters? *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: