

1. Title	Proposal for extensions to the Arabic block
2. Author	Mohamed Naji Grupellostr. 28 D-40210 Duesseldorf Gemany Mobile-Phone ::+49(0)152/18486011 Email : mohamednaji333@gmail.com
3. Submission date	the 6th of October, 2022

Sir/Madam

my proposals consists of three parts. This proposal presents several extensions to the Arabic script that are not currently included in the UCS repertoire.

I) part 1

with The Unicode Standard Range : 0600-06FF (we call it as UniversalForm) for the Arabic we can write Arabic text. However when we want to write some special like scientific text we will get difficulties.

If we want to write the physical unit millinewton meter [mNm] in Arabic, we have to write it like this :

- 1) without spaces
- 2) the first m must be connected with the N
- 3) the N must not be connected with the second m

we can only realize this, if

- 1) the first m must be in UniversalForm
- 2) the N must be in final form, so that it is not connected to the second m.
- 3) the second m must be in UniversalForm.
- 4) Between the first m and the N there must be a Tatweel (unicode 0640), so that the first m and the N are connected.

thus we will write

منم

it means (unicode 0645) + (unicode 0640) + (unicode FEE6) + (unicode FEE2)

we notice that to connect a glyph in UniversalForm with a glyph in final form, we have to interpose a Tatweel (unicode 0640), which is complicated. Therefore we propose to create a new form (let's call it closing form) for Arabic, so that any glyph in UniversalForm can be directly connected to a glyph in this new form (without Tatweel (unicode 0640)). This new form should satisfy further properties (see below)

Not only in physics but in many other fields such as programming, chemistry, mathematics and economics we write Arabic words one after the other without connecting them. The reader can usually hardly distinguish between such single words. For this we give the following example.

تفاجأ بأبيه

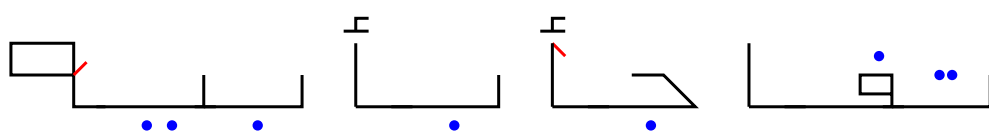
In the above example, we are dealing with two Arabic words that are placed one after the other. The two Arabic words are the Arabic word

تفاجأ

and the Arabic word

بأبيه

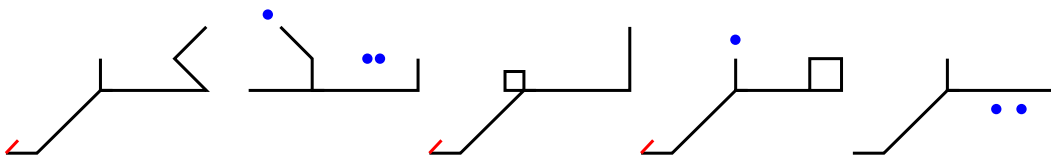
To distinguish the single words from each other, we suggest that the glyph be marked in closing form. For example, by a small check mark at the end, like this (the small check mark is in red) :



when we write

يرمز لوتذکر

we get difficulties to read it, because we can not distinguish between the single words. However when we use small check mark at the end it will be all clear



As a summary : we would like to create a new form (let's call it closing form) for arabic glyph so that

- 1) the glyph in closing form can be directly connected to a glyph in UniversalForm, if the glyph in UniversalForm is before the glyph in closing form
- 2) the glyph in closing form can not be connected to a glyph, which is after it
- 3) each glyph in closing form is marked so that it can be distinguished from any other form.

The glyph in closing form have exactly the same properties as the glyph in final form with the exception that it is directly (without Tatweel) connected to a preceding glyph in UniversalForm the glyph in closing form looks exactly the same as the glyph in isolated form with the only difference, that it the contains a small check mark at the end (in red)

We give now the representative glyphs and their suggested names :

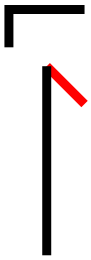


LETTER
ALEF WASLA CLOSING
FORM

LETTER
HAMZA
CLOSING
FORM



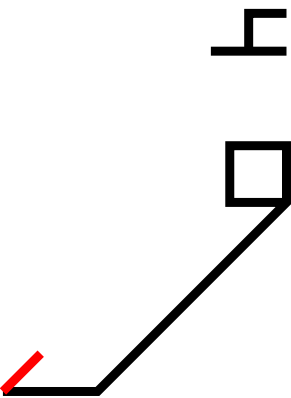
LETTER
ALEF
WITH
MADDA ABOVE
CLOSING
FORM



LETTER
ALEF
WITH
HAMZA ABOVE
CLOSING
FORM

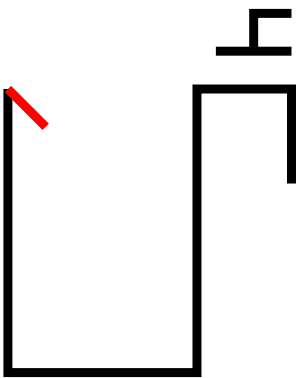


LETTER
WAW
WITH
HAMZA ABOVE
CLOSING
FORM





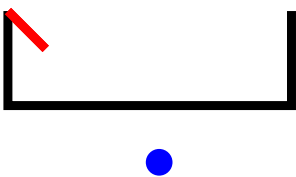
LETTER
ALEF
WITH
HAMZA BELOW
CLOSING
FORM



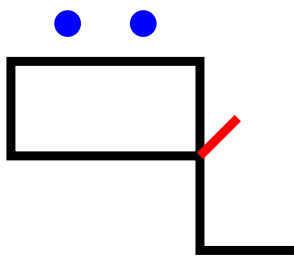
LETTER
YEH
WITH
HAMZA ABOVE
CLOSING
FORM



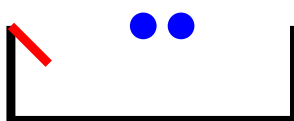
LETTER
ALEF CLOSING
FORM



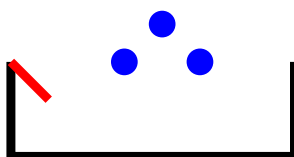
LETTER
BEH CLOSING
FORM



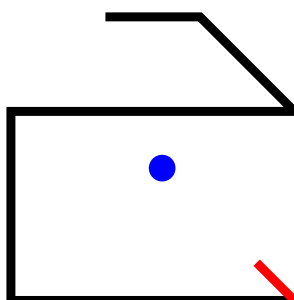
LETTER
TEH MARBUTA
CLOSING
FORM



LETTER
TEH CLOSING
FORM

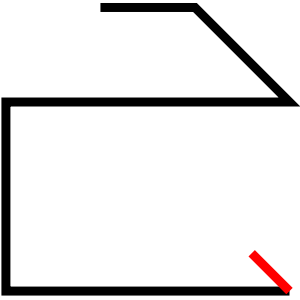


LETTER
THEH CLOSING
FORM

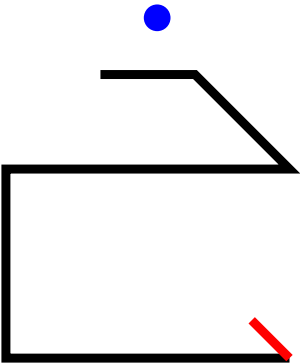


LETTER
JEEM CLOSING
FORM

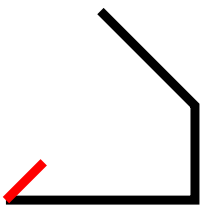
LETTER
HAH CLOSING
FORM



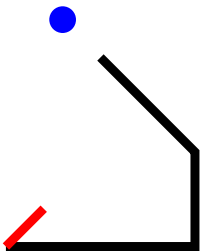
LETTER
KHAH CLOSING
FORM



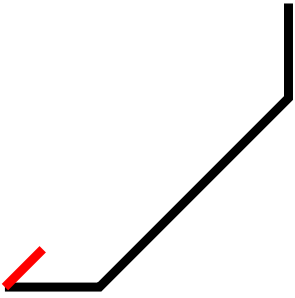
LETTER
DAL CLOSING
FORM



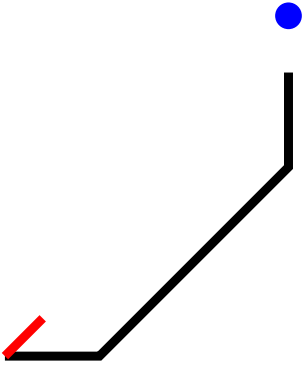
LETTER
THAL CLOSING
FORM



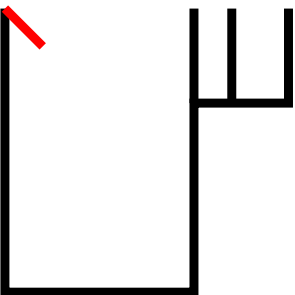
LETTER
REH CLOSING
FORM

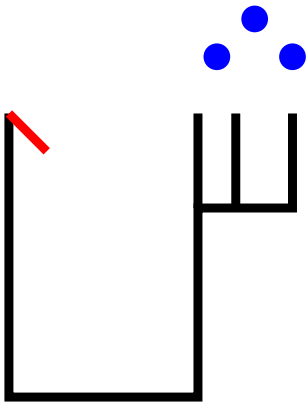


LETTER
ZAIN CLOSING
FORM

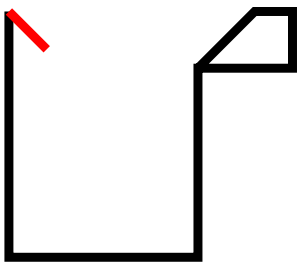


LETTER
SEEN CLOSING
FORM

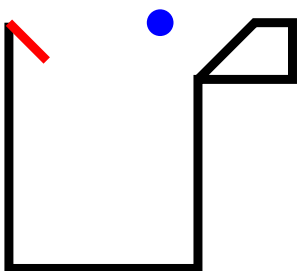




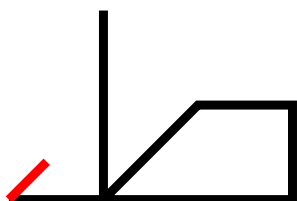
LETTER
SHEEN CLOSING
FORM



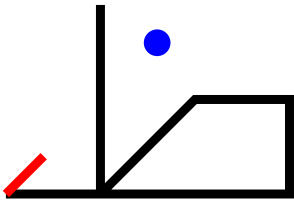
LETTER
SAD CLOSING
FORM



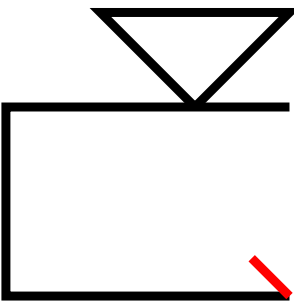
LETTER
DAD CLOSING
FORM



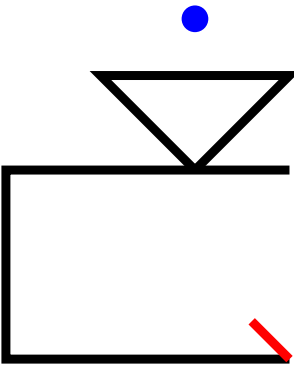
LETTER
TAH CLOSING
FORM



LETTER
ZAH CLOSING
FORM



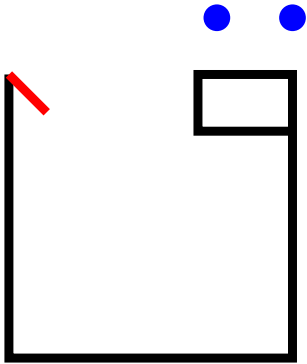
LETTER
AIN CLOSING
FORM



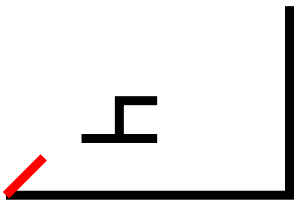
LETTER
GHAIN CLOSING
FORM



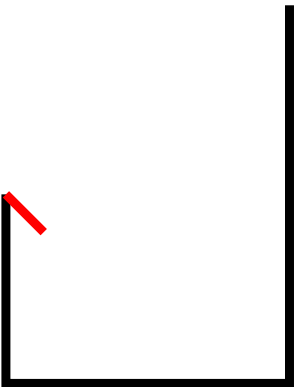
LETTER
FEH CLOSING
FORM



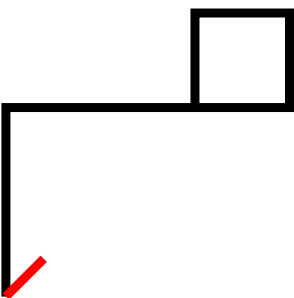
LETTER
QAF CLOSING
FORM



LETTER
KAF CLOSING
FORM

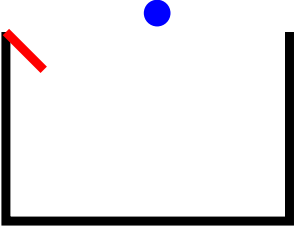


LETTER
LAM CLOSING
FORM

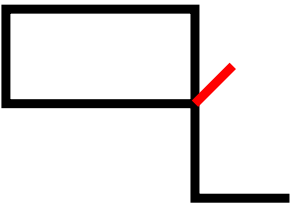


LETTER
MEEM CLOSING
FORM

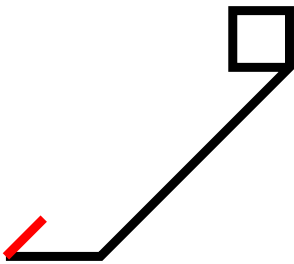
LETTER
NOON CLOSING
FORM



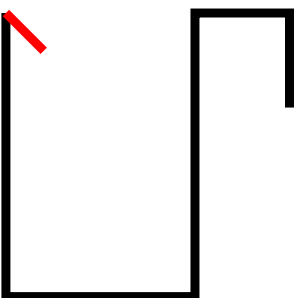
LETTER
HEH CLOSING
FORM



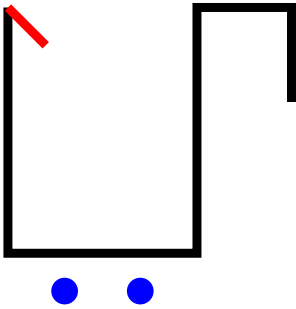
LETTER
WAW CLOSING
FORM



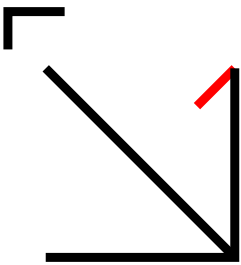
LETTER
ALEF MAKSURA CLOSING
FORM



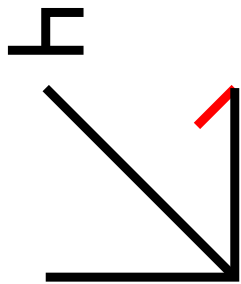
LETTER
YEH CLOSING
FORM



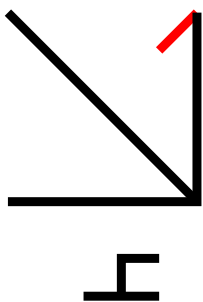
LIGATURE LAM
WITH
ALEF WITH
MADDA ABOVE
CLOSING
FORM

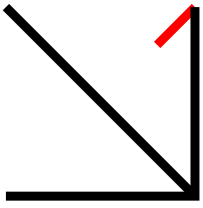


LIGATURE LAM
WITH
ALEF WITH
HAMZA ABOVE
CLOSING
FORM

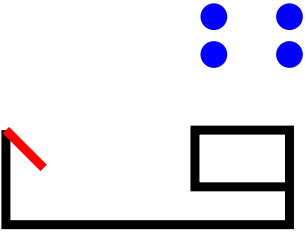


LIGATURE LAM
WITH
ALEF WITH
HAMZA BELOW
CLOSING
FORM

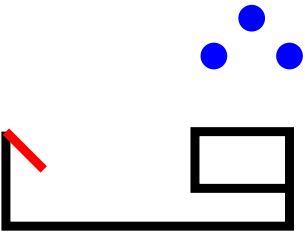




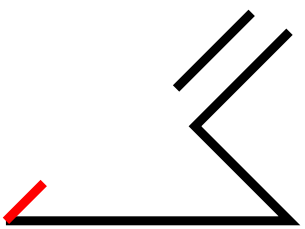
LIGATURE
LAM WITH
ALEF CLOSING
FORM



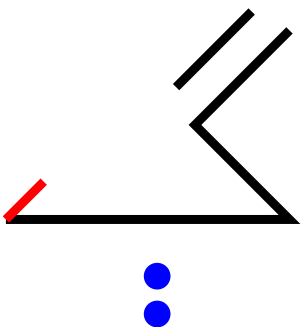
LETTER
PEHEH CLOSING
FORM



LETTER
VEH CLOSING
FORM

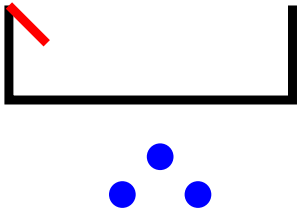


LETTER
GAF CLOSING
FORM

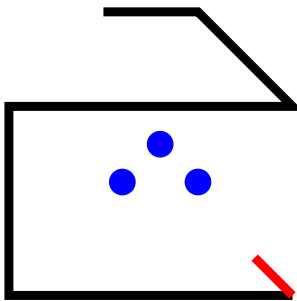


LETTER
GUEH CLOSING
FORM

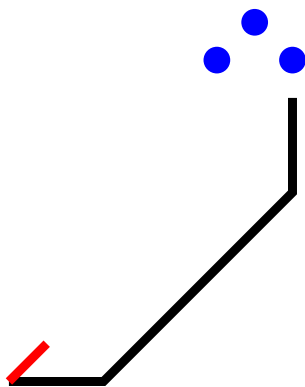
LETTER
PEH CLOSING
FORM



LETTER
TCHEH CLOSING
FORM



LETTER
JEH CLOSING
FORM



II) part 2

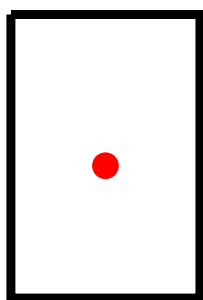
The Characters Digit Zero to Digit Nine (The Unicode Standard Range : 0030-0039) cause problems in the Arabic text (specially in a scientific text) because they have European Number as Bidirectional Class and are not Mirrored. But many Arabic countries use these numerals and neither the Arabic-Indic numerals (The Unicode Standard Range : 0660-0669) nor the Extended

Arabic-Indic numerals (The Unicode Standard Range : 06F0-06F9).

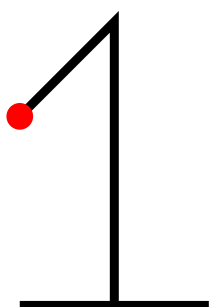
If we want to write a phone number with International Prefix and National Code, we get an incorrect order of the numbers, for example : when we write +216(0)123/456789 we will get for Arabic +612(0)987654/321

So I would like to add few new digits based on the idea of al-Khwarizmi
(https://en.wikipedia.org/wiki/Muhammad_ibn_Musa_al-Khwarizmi)

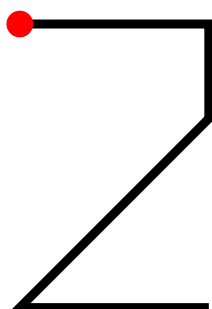
We give now the representative glyphs and their suggested names



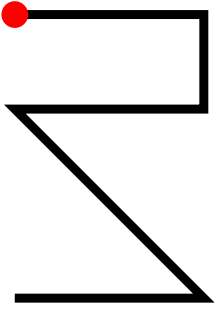
KHWARIZMI DIGIT ZERO



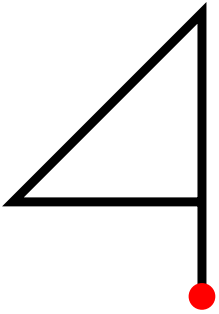
KHWARIZMI DIGIT ONE



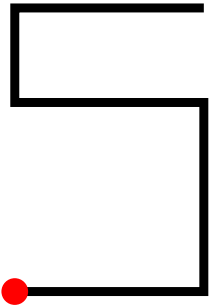
KHWARIZMI DIGIT TWO



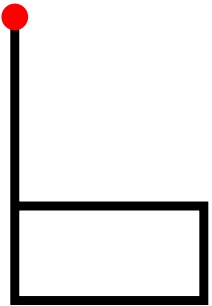
KHWARIZMI DIGIT THREE



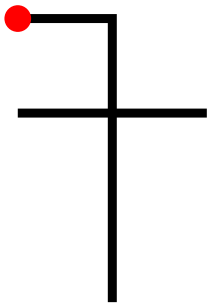
KHWARIZMI DIGIT FOUR



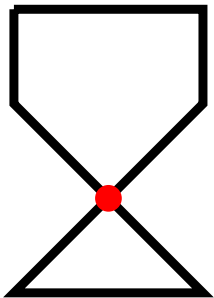
KHWARIZMI DIGIT FIVE



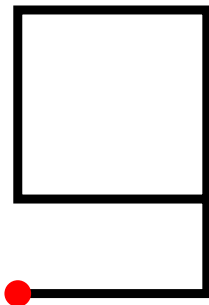
KHWARIZMI DIGIT SIX



KHWARIZMI DIGIT SEVEN



KHWARIZMI DIGIT EIGHT



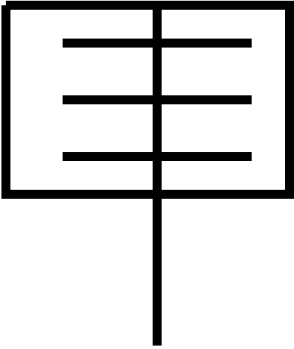
KHWARIZMI DIGIT NINE

As you see these Digits get a red point to distinguish them from the Digits with (The Unicode Standard Range : 0030-0039).

These digits must have Arabic as script and Arabic Number (AN) as Bidirectional Class.

III) part 3

finally we want to add the sign below, which was drawn by the Prophet Mohamed



SIGN DRAWN BY THE PROPHET MOHAMED

Thank you for your time !,

Sincerely
Mohamed NAJI