Arabic mathematical symbols in Unicode

Mohamed Jamal Eddine BENATIA*,
Azzeddine LAZREK** and Khalid SAMI***

*benatiamje@yahoo.fr, **lazrek@ucam.ac.ma, ***k_sami@ucam.ac.ma
Cadi Ayyad University, Faculty of Sciences
P.O. Box 2390, Marrakech, Morocco
Phone: +212 44 43 46 49 Fax: +212 44 43 74 09
http://www.ucam.ac.ma/fssm/rydarab

Abstract

The Unicode Standard provides a quite complete set of standard mathematic characters to support publication of mathematics in a Latin script based writing. Arabic alphabet based scripts make use of local ways for writing mathematics. Even tough some local symbols can be obtained via mirroring of already existing symbols, there are many symbols found in Arabic mathematical handbooks that are not yet part of the Unicode Standard and can't be obtained through a simple mirroring.

The paper presents a discussion of the usual mathematics character repertoire used in mathematics written in an Arabic presentation. Some of such special characters are submitted for inclusion into the Unicode Standard. This proposal also contains some non-mathematical characters such as units, arrows and dates symbols.

Key words: Arabic script, Arabic mathematical notation, Mathematical characters, Unicode.

1 Overview

There are two main ways to denote mathematics in Arabic script based languages. In some contexts, mathematical texts use the usual mathematical symbols just as they are in Latin script based texts. Mathematical expressions flow then from left to right against the stream of the natural language. In other contexts, mathematical texts use specific symbols spreading out from right to left in accordance with the natural language writing.

Actually, as the mathematical symbols used in Latin script based texts (see Table 1), there are *Arabic mathematical* symbols such as the mirrored symbol of *summation* ∑ or the conventional calligraphic symbol of *limit* نهـا.

U+2200-22FF	Mathematical Operators
U+2980-29FF	Miscellaneous Mathematical Symbols
U+2A00-2AFF	Supplemental Mathematical Operators
U+1D400-1D7FF	Mathematical Alphanumeric Symbols

Table 1: Main mathematical ranges in Unicode

These symbols are widely used in the writing of mathematics in languages written in an Arabic alphabet based script such as Arabic or Persian. Mathematical handbooks in Egypt, Syria, Libya, Iran, ... adopt generally the second option, especially in the K-12 education. Of course, in both the two mathematical notation systems, the mathematical expression has exactly the same meaning. Only the *way mathematical expressions are presented* is different.

Therefore, it should be quoted that there should exist a one to one mapping between the two mathematical notation systems, besides some differences according to the local area. So, in contrast with expressions in natural languages, mathematical expressions should be (and could be) *automatically translatable*. Can we build converters between different systems of local usages of mathematical symbols? The RyDArab system [3] [2] can do that up to some extent in a TEX environment. The Dadzilla system [9], an adapted version of Mozilla used for Arabic mathematics in MathML, can help in composing Arabic mathematical e-documents.

Of course, the use of a mathematical symbol can be localized. There can exist different local options to denote symbols. In fact, the symbols "," and "." denote the limit between thousands and units in decimal numbers respectively in English (e.g., 3.141 and 3,141). Conversely, in French, they are used in the opposite way. In the Arabic alphabet based mathematical notation, there can also exist different ways to denote the same concept. Both the symbols مجـ and $\sum$ can denote the operator *summation* and both the notations ٣,١٤١ and 3,141 represent the same decimal number.

All the proposed symbols contained in this paper come from Arabic handbooks of mathematics [7]. Most of these symbols had been adopted through official international conventions such as the *Amman's* 1987 *convention* [6] [11]. As such usages seem to be widespread, these symbols may be proposed to be added to the Unicode Standard [5].

2 Mathematical alphabetical symbols

Mathematical notation uses a basic set of mathematical alphanumeric characters consisting basically on the sets:

- Arabic digits: 0 - 9;
- Arabic-Indic digits: ٠ - ٩;

- Arabic letters: ا - غ ;
- specific Arabic letters: ﭫ, ... ;
- uppercase and lowercase Latin letters: A - Z and a - z;
- uppercase and lowercase Greek letters: A - Ω and α - ω;
- plus the other symbols.

For mathematical variables, style variations are meaningful [1]. Any style variation implies changes in the meaning of a given mathematical symbol. Even though Unicode encodes characters not glyphs, mathematical alphabetic symbols have many styles that differs from each other only by glyph. For Latin alphabet based Mathematical Alphanumeric Symbols, range 1D400-1D7FF, there are: Bold, Italic, Bold Italic, Script, Bold Script, Double-struck, Double Fracture, Sans-serif, Sans-serif bold, Sans-serif italic, Sans-serif bold italic and Monospace. For Greek, there are: Bold, Italic, Bold Italic, Sans-serif, Sans-serif bold, Sans-serif italic and Sans-serif bold italic.

Mathematical alphabetic symbols, in Latin based script texts, are typeset in a different way from that in ordinary text. For example, the character spacing is different in math mode than in text mode (e.g., let n be a number in the set N subset of the natural numbers set $\mathbb{N}$ in the context $\mathcal{N}$). In right-to-left Arabic mathematical expressions, the usual shaping (ligaturing) of Arabic letters is omitted in math mode except for abbreviations or units such as trigonometric functions names.

The alphabetic mathematical letter-like symbols used in Arabic mathematical handbooks are of six forms: *isolated*, *initial*, *tailed*, *stretched*, *looped*, and *double-struck* (see Table 2). It should be noted that the tailed form is not contained in the Amman's convention but is commonly present in handbooks.

There are two alphabetic orders in Arabic. The one used in mathematics or numeration lists is the *a, b, j, d, ...* (ا, ب, ج , د, ...) order. It differs from the *a, b, t, th, ...* (ا, ب, ت, ث, ...) order in use in modern dictionaries. The order *a-b-j-d*[1] is the oldest one.

2.1 Basic mathematical alphabetic symbols

In order to avoid ambiguities, the Arabic character types are generally based on dot-less letters. As some Arabic letters differ only by the addition of dots below or above basic symbols, the basic *dot-less* symbols list is smaller than the complete list of the alphabet. On the other hand, the ambiguous dot-less letter-like symbols (see Table 3) should be named with care.

[1] The word *abjad* is not of Arabic origin, it comes from earlier written alphabets, perhaps from Phoenician or even from Ugaritic. In any way, it certainly predates the writing of Arabic, as can be seen by comparison of the Hebrew (Aleph, Beth, Gimel, Daleth, ...) and the Greek (Alpha, Beta, Gamma, Delta, ...) orders. In Gematrical numerical system, a number is assigned to each letter [10].

ج	MATHEMATICAL JEEM ISOLATED FORM ≈ <font> <isolated> 062C ج Arabic letter jeem
جـ	MATHEMATICAL JEEM INITIAL FORM ≈ <font> <initial> 062C ج Arabic letter jeem ≈ <font> FE9F جـ Arabic letter jeem initial form
ـجـ	MATHEMATICAL JEEM TAILED FORM ≈ <font> FE9F جـ 06C1 ـہ
جا	MATHEMATICAL JEEM STRETCHED FORM ≈ <font> FE9F جـ FE8E ـا
ڿ	MATHEMATICAL JEEM LOOPED FORM ≈ <font> <isolated> 062C ج Arabic letter jeem
ج	MATHEMATICAL JEEM DOUBLE-STRUCK FORM ≈ <font> <isolated> 062C ج Arabic letter jeem

Table 2: Example of mathematical alphabetic symbols with and without dots

Letter	Pronunciation
ب د	BEH, because it comes before the TEH and THEH
ح خ	JEEM, because it comes before the HAH and KHAH
ر	REH, though it comes after the ZAIN
ف ۋ	FEH, because it comes before the QAF
ق	QAF, because it hasn't a glyph similar to the FEH in isolated form

Table 3: dot-less letter-like symbols ambiguous names

In order to provide a big amount of mathematical symbols and to satisfy either the needs of those who use dot-less characters and those who prefer characters with dots, alphabetic mathematical symbols may be proposed with and without dots. So, the proposition will remain in the Unicode Standard spirit who recommends representing characters not glyphs. Actually, the symbols are presented with and without dots in the *a-b-j-d* order (see Table 4).

The glyphs of the letters ALEF[2], HEH and KAF (ا, ه and ك respectively), in isolated and double-struck forms, can lead to some confusions. They don't appear in these forms in dot-less styles.

The glyphs of the letters ALEF, DAL, WAW, ZAIN, TAH, REH, THAL and ZAH (ا, د, و, ز, ط, ر, ذ and ظ respectively), in isolated form, are the same as in initial form. So, these symbols will appear in the isolated form set only.

[2]All along this paper, Arabic characters are named according to the Unicode Standard way, in spite of the non conformity of this appellation for some letters (e.g., the letter ز is generally pronounced ZAY instead of ZAIN).

The glyphs of the letters ALEF, DAL, WAW, ZAIN, REH and THAL (ا, د, و, ز, ر and ذ respectively), in the tailed and stretched forms, are composed with two elements. They won't appear in these forms in dot-less styles.

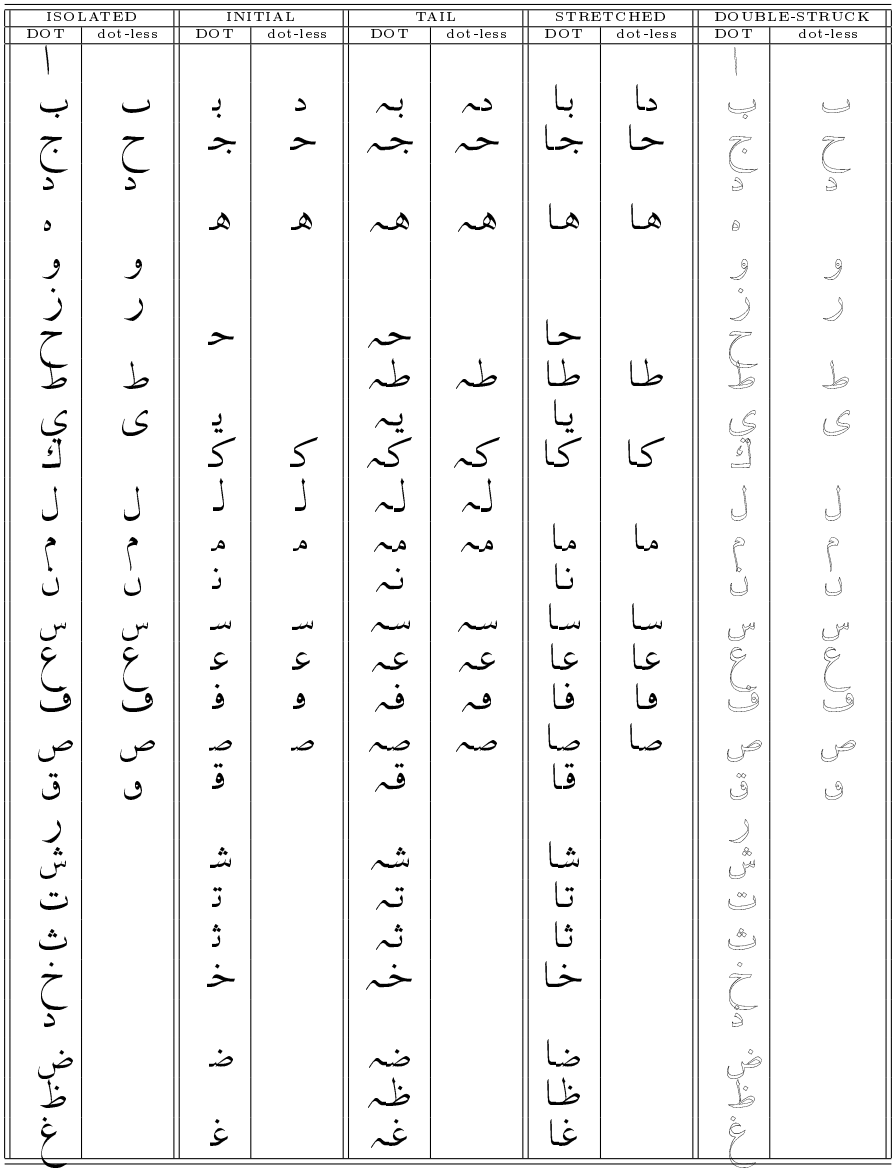

Table 4: Mathematical alphabetic symbols

2.2 Particular mathematical alphabetic symbols

Some glyphs used in Arabic mathematical presentation are not really Arabic letters but particular forms of mathematical alphabetic symbols used in Arabic mathematical handbooks (see Table 5 and 6).

- The glyph of the letter ALEF ا can be taken for the Arabic-Indic digit ONE ١. Thus, it is replaced by أ.
- The glyph of the letter HEH ه can be taken for the Arabic-Indic digit FIVE ٥ in the isolated and double-struck forms. Thus, it is replaced by هـ.
- The glyph of the letter KAF ك results from the combination of two elements in the isolated and double-struck forms. Thus, it is replaced by either ك or ک.
- The glyph of the letter NOON can be found in different orientations and styles, with and without dot, (ن ں, ن ں, ن ں, ن ں, ن ں) according to the local area.

2.3 Exceptional mathematical alphabetic symbols

Some glyphs used in Arabic mathematical presentation are not really Arabic letters but symbols used in physics or in some Arabic alphabet based scripts such as Persian (see Table 7).

2.4 Large mathematical alphabetic symbols

The Arabic n-ary summation operator is denoted by either Σ and مجـ symbols according to the local area (see Table 8).

The Arabic n-ary product operator is denoted by either ∏ and جذ symbols according to the local area (see Table 8).

The Arabic limit operator is denoted by the symbol نها (see Table 8).

The Arabic factorial operator is denoted by either ! and ـا symbols according to the local area (see Table 8).

The symbols Summation, Product and Limit can be denoted with dots (مجـ, جذ and نها) or without dots (محـ, حد and ٮها).

Obviously, some software tools, such as TEX or MathML, can help to combine any text string with any symbols as needed. WG 2 adopted in Resolution M38.12 not to add any more Arabic presentation forms to the standard and suggests users to make use of appropriate input methods, rendering and font technologies that meet the user requirements. We propose these *large operators* for addition to the Unicode Standard even though they don't have similar entities in Latin. Actually, the shape of these ligatures is unusual compared to the layout in

أ	MATHEMATICAL ALEF ≈ <font> <isolated> 0627 ا Arabic letter alef
،	MATHEMATICAL DAL ≈ <font> <isolated> 062F د Arabic letter dal
ھ	MATHEMATICAL HEH ≈ <font> <isolated> 0647 ه Arabic letter heh
ے	MATHEMATICAL YEH ~~BARREE~~ [2] ≈ <font> <isolated> 06D2 ے Arabic letter yeh barree
ك	MATHEMATICAL KAF ≈ <font> <isolated> 0643 ك Arabic letter kaf
ک	MATHEMATICAL SWASH KAF ≈ <font> <isolated> 06AA ک Arabic letter swash kaf
لا	MATHEMATICAL LAMALEF ≈ <font> <isolated> FEFB لا Arabic ligature lam with alef
م	MATHEMATICAL MEEM ≈ <font> <isolated> 0645 م Arabic letter meem
ں ن	MATHEMATICAL ~~INVERTED~~ [2]NOON ≈ <font> <isolated> 0646 ن Arabic letter noon
ر	MATHEMATICAL REH ≈ <font> <isolated> 0631 ر Arabic letter reh
ز	MATHEMATICAL ZAIN ≈ <font> <isolated> 0632 ز Arabic letter zain
ء	MATHEMATICAL HAMZA ≈ <font> <isolated> 0621 ء Arabic letter hamza

Table 5: Particular forms of mathematical alphabetic symbols
[1]We strongly propose not to make use of this attribute.

regular text. Moreover, the size of these ligatures symbols varies according to the covered expressions (see Figure 1).

$$\sum_{ب=ط-ف-ح-1}^{ص}مج \qquad \sum_{ب=ط-ف-1}^{ص}مج \qquad \sum_{ب=ط-1}^{ص}مج$$

Figure 1: Variable-sized conventional summation operator

The n-ary operators like summation and integration may expand in size to

𝔸	MATHEMATICAL DOUBLE-STRUCK ALEF ≈ <font> <isolated> 0627 ا Arabic letter alef
𝔻	MATHEMATICAL DOUBLE-STRUCK DAL ≈ <font> <isolated> 062F د Arabic letter dal
ℍ	MATHEMATICAL DOUBLE-STRUCK HEH ≈ <font> <isolated> 0647 ه Arabic letter heh
𝕖	MATHEMATICAL DOUBLE-STRUCK YEH ~~BARREE~~ [2] ≈ <font> <isolated> 06D2 ے Arabic letter yeh barree
𝕜	MATHEMATICAL DOUBLE-STRUCK KAF ≈ <font> <isolated> 0643 ك Arabic letter kaf
𝕜	MATHEMATICAL DOUBLE-STRUCK SWASH KAF ≈ <font> <isolated> 06AA ک Arabic letter swash kaf
𝕝	MATHEMATICAL DOUBLE-STRUCK LAMALEF ≈ <font> <isolated> FEFB ﻻ Arabic ligature lam with alef
𝕞	MATHEMATICAL DOUBLE-STRUCK MEEM ≈ <font> <isolated> 0645 م Arabic letter meem
∩ ∩	MATHEMATICAL DOUBLE-STRUCK ~~INVERTED~~ [2]NOON ≈ <font> <isolated> 0646 ن Arabic letter noon
𝕣	MATHEMATICAL DOUBLE-STRUCK REH ≈ <font> <isolated> 0631 ر Arabic letter reh
𝕫	MATHEMATICAL DOUBLE-STRUCK ZAIN ≈ <font> <isolated> 0632 ز Arabic letter zain
𝕩	MATHEMATICAL DOUBLE-STRUCK HAMZA ≈ <font> <isolated> 0621 ء Arabic letter hamza

Table 6: Double-struck particular forms of mathematical alphabetic symbols
[1]We strongly propose not to make use of this attribute.

fit with the expression they cover. The stretching can be performed by some software such as CurExt [4] [8]. These operators can have limits. As in the Latin alphabet based notation, the place of the limits in an operator is not the same in text expression inserted among words as in displayed expression alone in a line.

2.5 Combined symbols

The "equal to by definition" $\overset{\text{def}}{=}$ operator symbol exists already (see Table 9). An "equivalent to by definition" $\overset{\text{def}}{\Longleftrightarrow}$ operator symbol may be proposed for inclusion. A "combining definition" $\overset{\text{def}}{\bigcirc}$ may also be proposed, so, it can be

ٻ	MATHEMATICAL PEH ISOLATED FORM ≈ <font> <isolated> 067E ٻ Arabic letter peh
ٻ	MATHEMATICAL PEH LOOPED FORM ≈ <font> <isolated> 067E ٻ Arabic letter peh
ڃ	MATHEMATICAL TCHEH ISOLATED FORM ≈ <font> <isolated> 0686 ڃ Arabic letter tcheh
ڃ	MATHEMATICAL TCHEH LOOPED FORM ≈ <font> <isolated> 0686 ڃ Arabic letter tcheh
ڤ	MATHEMATICAL VEH ISOLATED FORM ≈ <font> <isolated> 06A4 ڤ Arabic letter veh
ڤ	MATHEMATICAL VEH LOOPED FORM ≈ <font> <isolated> 06A4 ڤ Arabic letter veh
ۼ	MATHEMATICAL AIN WITH TREE DOTS ABOVE ISOLATED FORM ≈ <font> <isolated> 06A0 ۼ Arabic letter ain with tree dots above
ۼ	MATHEMATICAL AIN WITH TREE DOTS ABOVE LOOPED FORM ≈ <font> <isolated> 06A0 ۼ Arabic letter ain with tree dots above

Table 7: Exceptional forms of mathematical alphabetic symbols

مجـ	محـ	ARABIC LARGE[2]N-ARY SUMMATION ≈ <font> FCCE مجـ Arabic ligature meem with jeem initial form
جذ	حد	ARABIC LARGE N-ARY PRODUCT ≈ <isolated> 062C ج 0630 ذ
نهـا	ههـا	ARABIC LARGE LIMIT ≈ <isolated> FCEF نهـ FE8E ا
لـ		ARABIC LARGE FACTORIAL ≈ FEDF لـ Arabic letter lam initial form

Table 8: Mathematical large symbols
[1]We propose to add the *large*, *alphabetic* or, at least, *conventional* adjective.

used with another character (see Table 10).

$\stackrel{\text{def}}{=}$	EQUAL TO BY DEFINITION

Table 9: Existing by definition symbol

$\overset{\text{def}}{\bigcirc}$	COMBINING DEFINITION
$\overset{\text{def}}{\Longleftrightarrow}$	EQUIVALENT TO BY DEFINITION

Table 10: Proposal by definition symbols

The Arabic operators "equal to by definition" and "equivalent to by definition" may be proposed as either compact symbols and element that can be combined with other symbols with dots or without dots (see Table 11).

تعر◌	عر◌	COMBINING ARABIC DEFINITION → xxxx $\overset{\text{def}}{\bigcirc}$
تعر$\overset{\text{def}}{=}$	عر$\overset{\text{def}}{=}$	ARABIC EQUAL TO BY DEFINITION → 225D $\overset{\text{def}}{=}$
تعر$\overset{\text{def}}{\Longleftrightarrow}$	عر$\overset{\text{def}}{\Longleftrightarrow}$	ARABIC EQUIVALENT TO BY DEFINITION → xxxx $\overset{\text{def}}{\Longleftrightarrow}$

Table 11: Arabic by definition symbols

Accented characters As is in Latin script based mathematical notation, only unaccented forms of the letters are used in Arabic mathematical notation. Mathematical symbols with diacritics are generally represented by combining characters sequences. Combining marks are used with mathematical alphabetic characters, instead of ready for use characters.

Mathematical functions names In general, basic Arabic characters, not combined with other signs, are used to represent single-character Arabic variables. In contrast, mathematical function names like jib (Sine), jibtamam (Cosine), mamass or zill (Tangent), etc., are represented by strings, ligatured characters, in usual text. This way allows distinguishing them from products of variables.

2.6 Units

The Arabic square unit, corresponding to the Latin square Km is marked with a special abbreviation كم. It represents an example of various signs for units that are worth of being included in the Unicode Standard (see Table 12).

2.7 Dates

Two main calendars are in use in the Islamic cultural area: the Hejry (Islamic calendar) and the MylAdy (Gregorian calendar). As the famous marks AC. and

كم ARABIC SQUARE KM
≈ <square>0643 كـ 0645 م
→ 339E square km

Table 12: Example of unit symbols

BC. used to distinguish years after and before the year zero. The marks هـ and م are used to make a distinction between the two calendars. Of course, these signs may be added as special characters (see Table 13).

هـ ARABIC HEJRY DATE
≈ <font> <isolated> 0647 هـ Arabic letter heh
م ARABIC MYLADY DATE
≈ <font> <isolated> 0645 م Arabic letter meem

Table 13: Calendar marks

3 Mathematical symbols

In the Unicode standard, mathematical symbols are named in reference with their:

- appellation (e.g., ∼ *tilde*);
- representation (e.g., ⟶ *right-arrow*);
- meaning (e.g., < *less-than*);
- representation and meaning (e.g., "(" *left parenthesis* = *open parenthesis*).

In Unicode, the official character name is frozen once it's published. This allows the use of the character names in identifiers, and this is useful for correlating Unicode and other standards.

Aspects of mathematical symbols localization The meaning of some mathematical symbols is not context free. For instance, in Arabic, the comma "،" as separator between items in a list is different from the one "," used to mark the limit between units and decimals in a decimal number. So, according to the context, the comma will have to be changed, into an Arabic comma or not.

The Arabic translation of the French or English number's list "1,2" is "2،1". In conterpart, the French decimal number "1,2", when translated to Arabic, remains unchanged. Moreover, in Arabic, the comma symbol glyph "،" is not

the mirrored form of the usual comma ",", so there is no issue here. This is down the translation process. The Unicode bidirectional algorithm is not relevant in this case - unlike, for example, summation or square root.

In Arabic calligraphy, the feather's "*kalam*" head is a flat rectangle. It is held so that the largest side is oriented so that it makes an angle of 70° with the baseline. Except for some variations, this orientation is kept all along the process of drawing the character [8]. As the Arabic writing goes from right to left, some boldness is produced around segments from top left toward the bottom right and conversely, segments from top right to the bottom left will rather be slim as in Figure 2.

Thus, some symbols would result from reversed glyphs, with respect to the vertical axis and to the horizontal axis go through the center of the glyph.

Figure 2: Sum symbol with its vertical and then horizontal mirrored image

In order to get mirrored symbols as used in Arabic mathematical expression, there are three possibilities:

- to eliminate the appellation in reference with the meaning. For instance, the symbol `0028` (becomes only *left parenthesis* and not also *open parenthesis*. Such eventuality is not allowed in the Unicode Standard where the name is frozen once introduced.
- to add the Arabic meaning appellation beside the old ones. For instance, the symbol `0028` (becomes *left parenthesis = open parenthesis* and also *= Arabic closing parenthesis*.
- to add those symbols. For instance, a new character `xxxx` (*left parenthesis = Arabic closing parenthesis*. The number of symbols will then be increased and there will be redundancy. The same symbol will have a double codification with two opposite meanings with respect to the context.

3.1 Mirrored symbols

Many symbols used in the Arabic presentation of mathematics are only horizontally reversed Latin mathematical symbols. For instance, the symbol ∋ *contains as member*, in a left-to-right writing denotes the meaning ∈ *element of* in a right-to-left writing and vice-versa. Most of these symbols, with their initial meaning, are already contained in the Unicode Standard. Concerning the names, there are three cases:

1. There is no clear relationship between names. For example, the symbol *element of* and its reversed form *contains as member*. It would be useful to add an equivalent new name (see Table 14).

2. The adjective `reversed` is added to the symbol's name (see Table 15). For example, the symbol *tilde* and its reversed form *reversed tilde*. Sometimes, some small changes are done.

3. The symbol's name contains the adjective `left` or `right` and its equivalent `open` or `closing`. For example, the symbol *left parenthesis* or *open parenthesis* and its reversed form *right parenthesis* or *closing parenthesis*. It may be useful to add an equivalent new name (see Table 16).

2208	∈	ELEMENT OF
		add = REVERSED CONTAINS AS MEMBER
		add ≈ <reversed> 220B ∋ contains as member
220B	∋	CONTAINS AS MEMBER
		add = REVERSED ELEMENT OF
		add ≈ <reversed> 2208 ∈ element of

Table 14: No relationship between names

223C	∼	TILDE OPERATOR
223D	∽	REVERSED TILDE
2243	≃	ASYMPTOTICALLY EQUAL TO
22CD	≌	REVERSED TILDE EQUAL

Table 15: Symbol and its reversed correspondent

0028	(	LEFT PARENTHESIS
		= OPEN PARENTHESIS
		add = REVERSED RIGHT PARENTHESIS
		add = ARABIC CLOSING PARENTHESIS
0029	)	RIGHT PARENTHESIS
		= CLOSING PARENTHESIS
		add = REVERSED LEFT PARENTHESIS
		add = ARABIC OPEN PARENTHESIS

Table 16: Left and its right symbols

The equivalence between `left` and `open` depends on the direction of writing. What opens in a left-to-right writing closes in a right-to-left writing and vice versa. For example, in a left-to-right context, `U+0028` is interpreted as open parenthesis and will appear as "(". While, in a right-to-left context, the same meaning will appear as the mirrored glyph ")". Then, the symbol is better described in reference with the meaning (ex. OPEN PARENTHESIS) instead

of in reference to the presentation (ex. LEFT PARENTHESIS)[3]. For instance, the expression $(a-2)$ in a left-to-right writing can be described as the sequence:

OPEN PARENTHESIS, a, $-$, 2, CLOSING PARENTHESIS,

or

LEFT PARENTHESIS, a,$-$, 2, RIGHT PARENTHESIS.

In a right-to-left context, for the equivalent expression $(2-\backslash)$ the description will rather be:

OPEN PARENTHESIS, $\backslash$, $-$, 2, CLOSING PARENTHESIS,

or

RIGHT PARENTHESIS, $\backslash$, $-$, 2, LEFT PARENTHESIS.

The words `right` and `left`, referring to the written presentation of the symbol, may so be changed into `preceding` and `following` when the reference is the meaning.

There are some frequently used Arabic symbols that have appropriate mirrored characters in the Unicode Standard (see Table 17). As, the number of those characters is not very important, many characters and their mirrored symbols (e.g., $<$ and $>$) are both already encoded with some ambiguity in the name, we propose them for inclusion into the Unicode Standard.

Other symbols are not frequently used. The corresponding regular character has the mirrored property, then, characters will be displayed with a mirrored image in right-to-left text composition.

Negated symbols The oblique bar in negated symbols is oriented in the Arabic mathematical symbols just as it is in the Latin ones. So, mirroring the negated symbol can sometimes lead to mistakes (see Table 18). Generally, the division symbol is the slash symbol (see Table 19). Once mirrored, it will become a backslash instead of the usual slash.

3.2 Not mirrored symbols

There are some Arabic symbols with no appropriate mirrored characters in the Unicode Standard (see Table 20). However, the available fonts and rendering engine could now provide the right mirrored glyphs of those symbols. We propose them to be added in the Unicode Standard.

4 Miscellaneous symbols

4.1 Particular symbols

In Arabic, there are particular symbols with no relation with the usual ones (see Table 21). They can be proposed for addition into the Unicode Standard.

[3]In Unicode version 1, the name of character (was OPENING PARENTHESIS. It has been changed, in later version, into LEFT PARENTHESIS = OPEN PARENTHESIS.

⅁	REVERSED COMPLEMENT ≈ <reversed> 2201 ∁
ⅆ	REVERSED PARTIAL DIFFERENTIAL ≈ <reversed> 2202 ∂
⅄	REVERSED THERE EXISTS ≈ <reversed> 2203 ∃
⅀	REVERSED N-ARY SUMMARY ≈ <reversed> 2211 ∑
⅋	REVERSED SQUARE ROOT ≈ <reversed> 221A √
⅂	REVERSED RIGHT ANGLE ≈ <reversed> 221F ∟
⅄	REVERSED ANGLE ≈ <reversed> 2220 ∠
⸧	REVERSED MEASURED ANGLE ≈ <reversed> 2221 ⦟
⸢	REVERSED SPHERICAL ANGLE ≈ <reversed> 2222 ∢
⅋	REVERSED INTEGRAL ≈ <reversed> 222B ∫
⅋⅋	REVERSED DOUBLE INTEGRAL ≈ <reversed> 222C ∬
⅋⅋⅋	REVERSED TRIPLE INTEGRAL ≈ <reversed> 222D ∭
⅋	REVERSED CONTOUR INTEGRAL ≈ <reversed> 222E ∮

Table 17: Frequently used symbols with appropriate mirrored image

2260	≠	NOT EQUAL TO ≡ 003D = 0338 ◌̸
	∄	REVERSED THERE DOES NOT EXISTS ≈ <reversed> 2204 ∄

Table 18: Negation symbols

4.2 Arrows

The signs listed in the range U+2790-27FF don't include symmetrical signs oriented right-to-left. So, supplemental arrows in Dingbat should be added. In particular, symmetrical signs for all symbols from 2794 to 27BE can be proposed with the specification LEFTWARDS, in contrast to RIGHTWARDS.

00BD 1/2 VULGAR FRACTION ONE HALF
• bar may be horizontal or slanted
≈ 0031 1 2044 / 0032 2

Table 19: One half symbol

REVERSED CUBE ROOT
≈ <reversed> 221B
REVERSED FOURTH ROOT
≈ <reversed> 221C
REVERSED ARABIC-INDIC CUBE ROOT
≈ <reversed> 221B
REVERSED ARABIC-INDIC FOURTH ROOT
≈ <reversed> 221C

Table 20: Mathematical symbols with no appropriate mirroring

Stars, asterisks and snowflakes
OUTLINED WHITE STAR
• Morocco sign

Miscellaneous symbols
CRESCENT

Letter-like symbols
ARABIC DIAMETER

General punctuation
ARABIC-INDIC PERMILLE SIGN
→ 2030 ‰ permille sign
ARABIC-INDIC PER TEN THOUSAND SIGN
→ 2031 ‱ per ten thousand sign

Table 21: Particular symbols

5 Conclusion

In Arabic alphabet based scripts, mathematical expressions are, in many contexts, written from right-to-left. Specific symbols are then in use. A list of such established symbols are probably to include into the Unicode Standard. Since neither the Unicode bidirectional algorithm nor the symbols mirroring process are enough to get these symbols. Some of alphabetic symbols, units, dates,

particular mathematical symbols and arrows have also been presented in this paper.

References

[1] Unicode Technical Report #25,
http://www.unicode.org/reports/tr25.

[2] Mustapha Eddahibi and Azzeddine Lazrek, *Arabic scientific document composition*, International Conference on Information Technology and Natural Sciences (ICITNS 2003, Amman, Jordan), 2003, 294–301, http://www.ucam.ac.ma/fssm/rydarab/doc/communic/ammsys.pdf.

[3] Azzeddine Lazrek, *A package for typesetting Arabic mathematical formulas*, Die TEXnische Komödie, DANTE e.V. **13. (2/2001)** (2001), 54–66, http://www.ucam.ac.ma/fssm/rydarab/doc/communic/dtk2_01.pdf.

[4] Azzeddine Lazrek, *CurExt, Typesetting variable-sized curved symbols*, EuroTEX'2003: 14th European TEX Conference (Brest, France), 2003, http://www.ucam.ac.ma/fssm/rydarab/doc/communic/curext.pdf, pp. 47–71.

[5] Arabic mathematical symbols for Unicode, http://www.ucam.ac.ma/fssm/rydarab/english/unicode.htm.

[6] Arabic mathematical symbols from Amman's 1987 convention, http://www.ucam.ac.ma/fssm/rydarab/doc/unicode/convention.zip.

[7] Arabic mathematical symbols from handbooks, http://www.ucam.ac.ma/fssm/rydarab/doc/unicode/handbook.zip.

[8] Mostafa Banouni, Mohamed Elyaakoubi and Azzeddine Lazrek, *Dynamic Arabic mathematical fonts*, LNCS **3130** (2004), 149–157, International Conference on TEX, XML and Digital Typography, TUG2004, Xanthi, Greece, http://www.springerlink.com/index/URHRT2EYKYHH1RPA.

[9] Mustapha Eddahibi, Azzeddine Lazrek and Khalid Sami, *Arabic mathematical e-documents*, LNCS **3130** (2004), 158–168, International Conference on TEX, XML and Digital Typography, TUG2004, Xanthi, Greece, http://www.springerlink.com/index/TNV9QBN36PK5NA9W.

[10] Gematrical numerical system,
http://www.submission.org/miracle/gematrical.html.

[11] Union of the Arab scientific linguistic groupings, *Scientific symbols and method of their use in Arabic language*, Amman, Jordan, in Arabic, 1987.