Issues of Rendering Arabic Mathematical Notation in Computer Software

Mona Alshehri

10 June 2014

Supervised by: Prof. Stephen M. Watt

Abstract

Typeface technology has become quite complex over the years. There have been several attempts to use Arabic calligraphic styles in computer typography. These proved to be useful, but they had their shortcomings and drawbacks. Computational time cost and lack of Arabic script documentation were the most crucial issues with that work. In a few studies, the accuracy of results obtained also was an issue. In this paper, we study and analyze Arabic calligraphy rules and the challenges that they introduce to Arabic mathematical notation. We take into account that Arabic script is written from right to left and that it is justified using stretching letters that are ruled by specific techniques. We review the difficulties of writing Arabic mathematical notation in computer software as well as attempted solutions to these problems.

1. Introduction

Early computer software was designed to support only languages represented by Latin fonts, such as English. As computer usage spread throughout the world, demand increased for software to be available in languages other than English. The demand was met gradually, as new computer fonts emerged. Fonts for the Arabic language were introduced; however, these fonts failed to capture all of the language's unique characteristics.

Arabic is the native language of the Middle East and North Africa. Arabic script has two features which make it distinctive with respect to encoding. First, Arabic text is written from right to left (RTL). Second, the shape of a single letter changes depending on its position, i.e. whether the letter is isolated, or occurs at the beginning, middle, or end of a word. Arabic script is used for languages besides Arabic, such as Farsi, Kurdish, Pashto, Turkic, and Maltese Arabic (which uses the Roman alphabet). Some of these languages augment the basic Arabic alphabet with additional special letters.

In order to typesetting Arabic correctly, software must be able to display text from right to left and make sure the letterforms are displayed correctly depending on their positions within words [1]. Arabic script elements can take different geometric shapes through their ability to rotate and interconnect. Thus, creating and using an Arabic font in support of the calligraphic rules of Arabic requires advanced techniques for varying letter forms according to position within words and lines, and relative to neighboring letters [2].

Mathematical expressions in Arabic use specific symbols, but are not spared from the calligraphic rules of the language. Such expressions also use dynamic cursive styling and the RTL writing direction. Unfortunately, many mathematics software packages fail to provide complete support for Arabic calligraphic rules and therefore render Arabic mathematical formulae incorrectly. The aim of this report is to discuss these problems.

1.1. Report Purpose

Arabic cultural communities are varied and have different styles of writing mathematical notation. Therefore, finding a standard method for Arabic scientific writing requires terminology and standards agreement on the rules for typesetting. Software development aimed at processing Arabic mathematical notation can help to achieve this standardization goal.

In this report, we describe the issues surrounding the representation of Arabic mathematical language in electronic document formats such as MathML. Our goal is not to propose solutions, but to describe important issues and provide basic information for implementations.

1.2. History of Displaying Mathematical Expressions on The Computer

Early computers were designed for scientific calculations based on limited input mathematical expressions. Expression editors used linear expression formats in order to simplify the translation process. For instance, the expression $\frac{2}{7-x^3}$ was written as *2/(7-x**3)* [3].

As computer applications increased in diversity, typesetting systems and expression editors were designed to print and render mathematical equations. One such system, TeX, has become the most popular among academic writers and publishers. TeX is used to format text and mathematical notation. The system's ambiguous linear syntax does not, however, permit mathematical expression evaluation. For instance, the integral formula $\int x^2\,dx$ is written in TeX as `\int x^2 dx`. The characters `dx` can be any one of the following: differential of `x`, variable d multiplied by variable x, or variable dx. Thus, TeX is used only to output mathematical expressions [3].

The ambiguous semantics of linear mathematical expression syntaxes such as that used in the TeX system led researchers to seek out a standard for mathematical expression representation. Such a standard would capture both the presentation information and the meaning of mathematical expressions, permitting them to be shared among different software applications For example, an expression could be cut from a rendering application (such as a Web browser) and pasted into a different application for calculation.

The mathematical expression standard research culminated in the development of the MathML standard. Founded by The World Wide Web Consortium (W3C), MathML supports the rendering of mathematical expressions without ambiguity. A markup language, MathML is used to edit and process mathematical content across applications such as computer algebra systems and print typesetting systems, as well as render the content in Web browsers [4]. Further, this standard can be expanded to support many natural languages.

2. Arabic Mathematical Language

The writing system used in the Arabic language is also used in languages such as Farsi, Kurdish, Pashto, Turkic, and Maltese Arabic. The distinguishing feature of Arabic is the right-to-left writing direction (see Figure 1). Also, each letter takes a different shape depending on its position in a word (i.e. at the beginning, middle, or end) or whether or not it is an isolated character on a line. The language of mathematics uses symbols in organized formulae to express ideas and concepts [5]; but, it also shares some features with natural language. In the next sections, we review the characteristics and basic rules of Arabic writing, as well as diagnose the issues surrounding the use of Arabic character sets in mathematical notation.

Input Letters	Rendering Letters	Output Word
م ل ا س →	س ا ل م ←	سالم

Figure (1): The Direction of Arabic Writing Goes from Right to Left (Input Word is Salem)

2.1.Basic Characteristics of Arabic Script

In Arabic script, the shape of a word is more important than the shape of its individual letters. Arabic calligraphers emphasize word shape and shape the letters accordingly in the words or even the entire sentence.

There are two aspects to learning the rules of Arabic calligraphy. First, intensive training in writing the different characters is fundamental to understanding Arabic calligraphy. Second, cognitive knowledge of the aesthetic philosophy of Arabic calligraphy is required in order to fully understand the secret of this art.

2.1.1. Character Shape Variations

The Arabic alphabet contains 28 letters, which are {ا ، ب ، ت ، ث ، ج ، ح ، خ ، د ، ذ ، ر ، ز ، س ، ش ، ص ، ض ، ط ، ظ ، ع ، غ ، ف ، ق ، ك ، ل ، م ، ن ، ه ، و ، ي}. Most Arabic letters change shape according to where they appear in a word. Within a sentence, each letter is in one of four positions: isolated, or at the beginning, middle, or end of a word .(see Figures 2 and 3) [6], [7].

The six letters {ا ، د ، ذ ، ر ، ز ، و} are exceptions to these rules, because they do not have a 'middle' position shape and they have two positions shapes, which are:

- Isolated and at the beginning are the same shape
- In the middle and at the end of the word are the same.

Figure (2): Letter Shape Variations Related to the Position in the Word (from [6], p. 34).

Figure (3): Arabic letters which connect from the right side only and have two positions (from [6], p. 34).

2.1.2. Direction of Writing

Arabic text is written from right to left (RTL), while Latin languages flow from left to right (LTR) (see Figure 1). Arabic text is called 'bidirectional text', because it flows from right to left and can include script that is read from left to right. For instance, the numbers in Arabic text go forward visually from left to right [8]. The number 156 is read starting from the digit 1; however, two-digit numbers are read from right to left (see Figure 4).

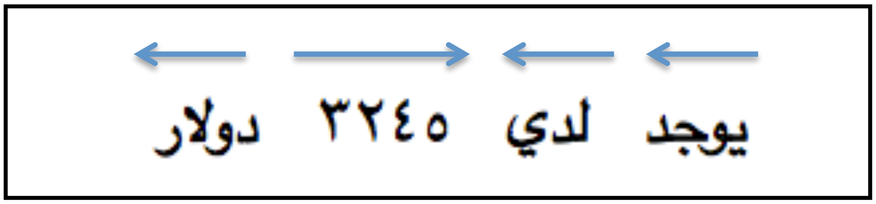


Figure (4): The direction of writing and reading in Arabic text is changed from RTL to LTR (like Latin languages) for numbers only.

The Unicode encoding supports two-direction text (bidirectional script) by storing the writing direction of characters within the characters themselves. That is, each character has a direction property. The Unicode Bidirectional Algorithm ('Bidi' algorithm) is used to manage the directionality of bidirectional Unicode text. For instance, the letters in right-to-left script are marked s RTL, and the letters in left-to-right script are marked as LTR [9].

2.1.3. Alternates

The different letter shapes in Arabic script are referred to as 'alternates'. This feature gives Arabic script a series of basic shapes (see Figure 5) [10]. Shape variations determine how one letter is connected to the next letter in a word, and provide the drawing style for each letter (which can change according to the writing purpose, e.g. note-taking vs. formal letter writing).

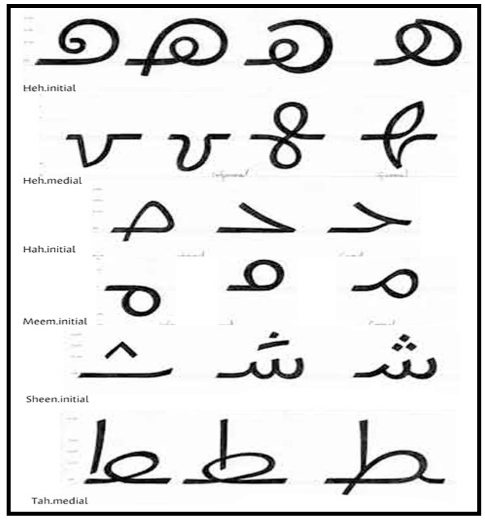


Figure (5) Variations for each shape
Source: Massira Font,WordPress,http://www.29arabicletters.com, Jan16/2013.

2.1.4. Linear Proportionality

Arabic calligraphy uses a linear proportionality rule. This rule uses character ratios to determine the relations between each character (the ratio of the length of a character to its width, for a given pen width). Also, characters are scaled by measuring length, width, depth, and slope. This scale is drawn from the top of the pen stroke and takes a rhombus or circle shape.

Researchers of Arabic calligraphy analyzed the forms of letters and found the *percentage of the virtuous*. This is a law followed by calligraphers for adjusting Arabic letters so as not to exceed the width and the length relative to other letters. We can define the percentage of the virtuous by using the first letter in the Arabic alphabet, Aleph, as a scale to which to compare the rest of the characters. For example, suppose a calligrapher writes Aleph at seven pen points in length. Assuming that the width of the Aleph is appropriately scaled to its seven-point length, a caliper is used to draw a circle around the Aleph to surround the entire letter, which is in the center of the circle (see Figure 6). This circle is the scale for the rest of the characters, which must not exceed the circle circumference [11].

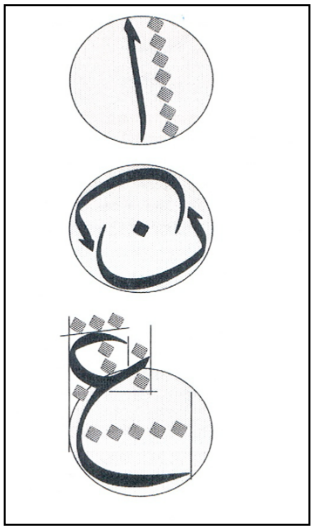


Figure (6) The proportional rule after drawing the circle around (Aleph). The body of the letters is inside the circle. The diameter of the circle is equal to the height of the letter Alef (from[10], p. 16).

Although some letters share the basic shape scale, Arabic text does have letters that differ only by parts of glyphs, or by the number and the position of dots. In Figure 7, the top letters share a part of a glyph. However, the number of dots and the position of them give rise to different letters. The lower letters share a part of a glyph, but adding new features to the basic form (such as a stretched or looped tail) produces new letters.

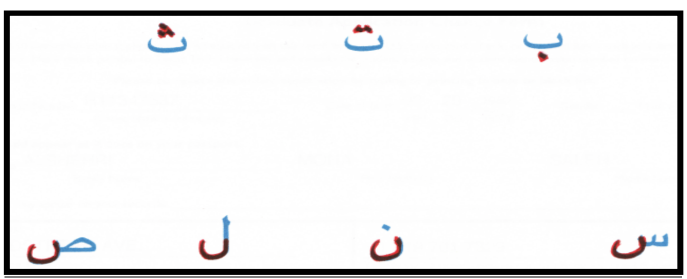


Figure (7) The upper letters share the same glyph, but differ by the number and position of dots. The lower letters share only part of the glyph.

2.1.5. Connected Letters

The natural flow of writing cursively is one of the most important characteristics of Arabic script. There are several techniques for efficiently connecting letters. First, character forms and rules are used to connect consecutive characters (see Figure 8) [6], [7]. Second, the position of the adjacent letter and cursivity rules are taken into account. A third method is to add consecutive words until a regular line is formed. Finally, Arabic script is justified by using horizontal elongation connections between the letters in words. This type of expansion in Arabic is called a *kashida*.

Figure (8) Transition from separate letters to joined characters in the word "Mohammed".

2.1.6. Kashidas

In Latin text, words are justified by using spacing between words and characters, and sometimes by hyphenation. In contrast, Arabic script is justified by using kashidas, which are simply elongations made to the connections between letters. Kashidas are slightly curved in shape, and there are rules for their formation (see Figure 9). Kashidas are used to fix crowded letters in a word, to emphasize a specific letter or word, and to create a balance in lines.

Most current typesetting software uses only limited technique for connecting Arabic characters (and omits support for dynamic characters). Typically, horizontal links are used between characters, and these are not consistent with Arabic calligraphy. However, kashidas would be easy to incorporate into predetermined typefaces, taking into account that they are not characters, but only elongations within words. Also, a kashida does not change when copied or when a word is cut and pasted into a different location [12].

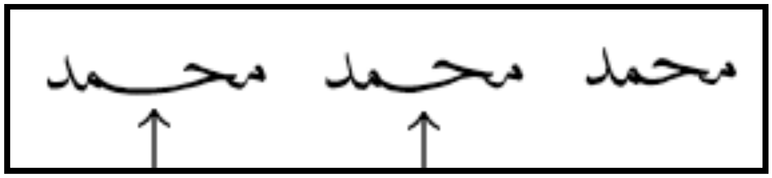


Figure (9): Various curvilinear kashidas.
Source: Mohamed Elyaakoubi, The Journal of Electronic Publish, Justify Just or Just Justify, http://dx.doi.org/10.3998/3336451.0013.105,Winter 2010

2.1.7. Vowels

In Arabic script there are two kinds of vowels. The first kind are the letter vowels (called *long vowels*), which are Alef, Waw, and Ya. These are used to show the stress on a given vowel. The second kind are the vowels marks (also called *short vowels* or *diacritic marks*), which are Fattha, Dhamma, Kasra, Sokoon, Shadda, and Maddah.

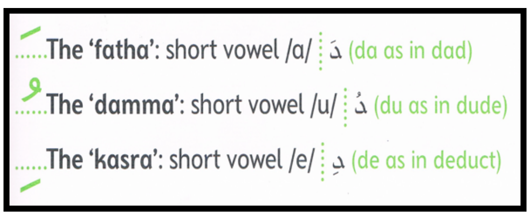
Diacritic marks are signs or symbols that are combined with the letters of the Arabic alphabet to indicate different pronunciations. Additionally, diacritic marks give Arabic script aesthetic form and fill in empty spaces. Each sign has a specific position on a character, which is determined based on several factors. The important factor is the elimination of ambiguity.

Diacritic marks are placed above or below letters in order to clarify their vocal pronunciation. Arabic script can be written without such vocalization marks; however, since the writing of the Quran it has become the norm to show how words are meant to be read. For example, k-t-b can be read in different ways if vocalization marks are omitted. It can be read as kataba ("he wrote") or as kutiba ("it was written"). This distinction is important, for example, in search engine use, since the correct meaning of the word cannot be established without vocalization marks [10].

Each of the five vocalizations (Fattha, Dhamma, Kasra, Sokoon, Maddah, and Shadda) has a different meaning for intended "mouth shape" (see Figure 10a):

- Fattha means "an opening"
- Dhamma means "a closing"
- Kasra means "a breaking"
- Sokoon means "Static (absent) or vowel-less"
- Shaddah means “double consonant” and it can be compared with Shadda, Fattah, Kasra and Dammah
- Maddah, the ”extension sign”, is used to extend the sound of an Arabic letter

In addition to the diacritic marks, there are marks called *tanween marks*. They are used to double short vowels to indicate the indefinite article (see Figure 10b).



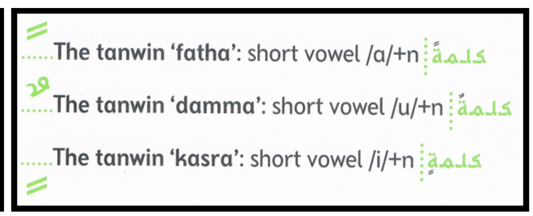


Figure (10a) **Figure (10b)**

Figure (10a) Vocalization Marks (Fattah, Dhamma, Kasrs) (from[6], p. 45).
Figure (10b) Tanween Marks (from[6], p. 47).

2.1.8. Diacritic Dots

Old Arabic calligraphy was written without dots. Dots were added in the ninth century to avoid confusion with the pronunciation of letters [6]. Up to three dots are added above or below a character form to create new letters, as in, for example, ت ، ب, ث.

2.1.9. Miniature Letters

In Arabic script, there is a small letter called *Hamza* that is added to the basic 28-letter Arabic alphabet. The reason for this additional letter (or sign) is to clarify minute letter shape variations (see Figure 11).

Figure (11) Hamza ' ء' in different positions on the letters.
Source: Submission, Submission.org,
http://submission.org/verify_writing_of_Alif_Yaa_and_Taa.html,2013.

2.1.11. Embellishment Marks

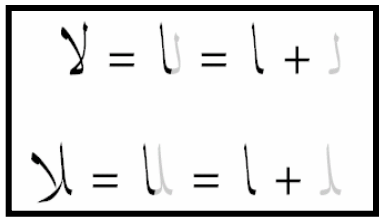
Embellishment marks are used to make Arabic script look pleasant. There is no specific reason to use this kind of mark other than to adorn text [13] (see Figure 12). Such marks are commonly used for names, titles, or poetry (see Figure 13).

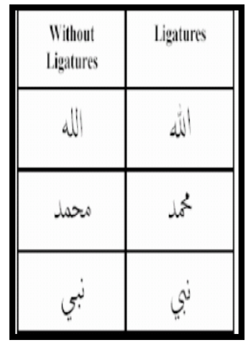
Figure (12) Embellishment Marks (from [10], p. 14)

Figure (13) Embellishment Marks Shapes (from [7])

2.1.12. Ligatures

Letters in Arabic script merge and change shape due to their cursive characteristics. For example, the letters 'lam ل' and 'alif ا' are merged together to form a new letter called 'lamalif' (see Figure 13a). In Latin script, an example similar to the 'lamalif' is the 'ampersand' character ('&'), which means 'and'. In Arabic calligraphy, ligatures serve to emphasize the quality of a word shape such that perfect combinations of individual letters are produced (see Figure 13b).





Without Ligatures	Ligatures
الله	الله
محمد	محمد
نبي	نبي

Figure (13a) **Figure (13b)**

Figure (13a): 'lamalif' is the required ligature that combines the letter 'lam' and 'alif '.
Source: http://quod.lib.umich.edu/j/jep/3336451.0013.105?rgn=main;view=fulltext
Figure (13b): Endless series of ligatures (from[6], p. 61).

2.2. Basic Characteristics of Arabic Mathematical Notation

Natural languages, such as English and Arabic, evolved in order to communicate human thought in the form of words. Mathematics, also a mental activity, is expressed in natural languages using not only words, but symbols and numbers as well. Real world problems are formulated as mathematical problems [5].

Mathematics is a language used to explain science alongside natural languages. It is used in daily life as well as in many disciplines, such as computer science, engineering, physics, and statistics. Mathematics can be defined as a group of related fields, including algebra, geometry, and calculus. It is the study of numbers, quantity, and shape and space and their inter-relationships, all expressed using specialized notation [5].

The rich knowledge base of science is now enhanced by publishing tools for mathematical notation, and the knowledge revolution has led to the distribution of scientific information across many channels: electronic documents, printed paper, audio and visual presentations, and via the World Wide Web. These tools do not, however, fully support Arabic mathematical notation. Although the Arabic mathematical language has remained the same for centuries, languages other than Arabic became the norm when modern mathematics emerged. It is for this reason that mathematical expressions are commonly written using different notation.

The absence of adequate scientific publishing tools for Arabic negatively effects Arabic education and scientific productivity. For instance, there are few Arabic scientific books that contain Arabic symbolic formulas and which are published in Arabic. The problem is further complicated by the fact that the cultures and regions where mathematics is used each have different notational styles; thus, improvements to typesetting tools would need to encompass all local Arabic mathematical styles [14]. We present more detail on localized styles in the following chapter.

In order for the Arab world to develop and keep pace with scientific progress, there is no choice but to teach science and technology in the Arabic language at all levels, including higher education. Therefore, rendering (as well as representing content in) Arabic scientific writing in computer software is much needed and can be accomplished through the following:

- Fixing, unifying, and integrating scientific terms that are used in Arabic countries
- Adjusting the words of the logical connections between mathematical formulas and notations
- Adopting standards and criteria for Arabic shapes and writing in the direction of the Arabic language
- Electronic academic references that are published in Arabic

The former points require the development of electronic publishing applications for producing high-quality, typeset documentation for which searching, indexing, and accessing content can be performed. The technical requirements for achieving localization of Arabic notation require an understanding of the basic characteristics of Arabic mathematical notation. The remainder of this chapter is devoted to describing the universal writing rules for the Arabic mathematical language.

In Arabic mathematical layout, expressions are characterized by the following properties.

2.2.1. Alphabetic Order

Arabic mathematics arose from a need to do business calculations, measure quantities and lengths, and to predict astronomical events. These beginnings led to the definition of three main aspects of mathematics: the studies of structure, space, and variables [15].

The alphabetic order used in Arabic mathematical notation differs from that of Arabic typescript. In the Arabic language, the arrangement of the mathematical alphabet follows a graphical system, where letters are sorted according to numeric value and grouped according to similar shapes (see Table 3) [6].

ي	و	ه		ن	م		ل	ك		ق	ف		غ	ع		ظ	ط		ض	ص	ش	س	ز	ر	ذ	د	ج	ح	خ		ث	ت	ب	ا
غ	ظ	ض		ذ	خ	ث		ت	ش	ر	ق		ص	ف	ع	س		ن	م	ل	ك		ي	ط	ح		ز	و	ه		د	ج	ب	ا
1000	900	800		700	600	500		400	300	200	100		90	80	70	60		50	40	30	20		10	9	8		7	6	5		4	3	2	1

Table (3): Arabic Letters in the graphical system (called *abjad*). The upper row represents the Arabic text letters, while the second row is the alphabetic order .The third row contains the numeric values.

2.2.2. Mathematical Letter Variations

The various shapes of Arabic letters are determined by position. A letter can occur in one of four different positions: in isolation, or at the beginning (initial), middle, or end of a word. Arabic mathematical notation introduces various additional letter styles in order to represent the enormous group of symbols that are used in mathematical expressions (see Figure 14) [16]. The Arabic mathematics letter styles are: tailed, looped, stretched, and double-struck, plus the 'initial' and 'isolated' styles.

	isolated	initial	tailed	looped	stretched	double-struck
dotted	𞸂	𞸢	𞹂	𞺂	𞹢	𞺢
undotted	𞸇	𞸧	𞹇	𞺇	𞹧	𞺧

Figure (14): Variation of Arabic Mathematical Symbols (from [2]).

2.2.3. Diacritic Dots

Alphabetic symbols in Arabic mathematical language are used with and without diacritic dots (see Figure 15). However, characters are typically formed by dot-less letters in Arabic mathematics notation (see Figure 16) in order to avoid ambiguity between the two types of alphabetic order (Arabic mathematical order and Arabic text order) [16]. Local language preferences also determine the use of dotted and un-dotted characters, taking into account that mixing the two styles within one document is not recommended. The dotted style is incorporated into symbol fonts in order to have all symbols available within Web browsers [3].

In Figure 16, the letters {ALEF, DAL, WAW, ZAIN, TAH, REH, THAL and ZAH} (ا ، د ، و ، ز ، ت ، ر ، ذ ، ز) have the same shapes in both isolated and initial form. Therefore, the set of isolated forms is used in the order for them. In addition, the letters {ALEF, DAL, WAW, ZAIN, REH, THAL} (ا ، د ، و ، ز ، ر ، ذ), in the tailed and stretched forms, are composed of two parts. So, they are not used in the dot-less style [16].

There are six letters in Arabic script, {و ، ز ، ر ، ذ ، د ، ا}, that do not have distinguished initial or middle position shapes. As a result, if these letters are used between two mathematical symbols, the expression appears as a product of two symbols. Therefore, to avoid confusion, these letters are not commonly used in the middle of mathematical expressions [2].

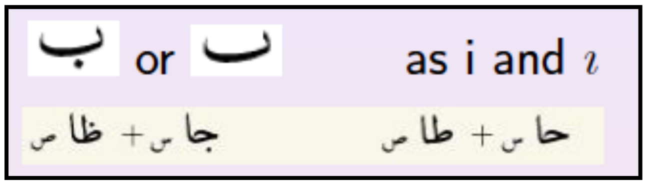


Figure (15): Arabic Mathematical Alphabetic Symbols are used with or without dots
Source: http://www.ucam.ac.ma/fssm/rydarab

Figure (16): The Arabic Mathematical Alphabetic Order (Dot and Dot-less Letters) (from [16])

2.2.4. Alphabetic Symbol Cursivity in Arabic Mathematical Expressions

Cursive Arabic writing is characterized by the use of rules for shaping letters based on their position within words and sentences. Within Arabic mathematical expressions, adjacent letters and symbols are written according to the following two rules:

- Alphabetic symbols are not written cursively

- Abbreviations are written cursively, i.e. with connected letters

We now explain in more detail these two different cursive modes used in Arabic mathematical writing.

2.2.4.1. Writing Non-Cursively

Unlike the cursive writing of alphabetic letters in Arabic script, mathematical expressions are written without connections between alphabetic letters. Symbols are used instead to convey the intended meaning (see Table 4) [15].

	Arabic	**English**
Script	ابس	abc
Mathematical	س٢ ب + ا	$A+bc^2$

Table (4): No Cursivity Between Adjacent Alphabetic Symbols in Arabic Mathematical Expressions

2.2.4.2. Writing Cursively

In Arabic notation, abbreviations (e.g. describing variables or units) using more than one letter are written as a single ligature. That is, abbreviations are written cursively, with connected letters. Unlike Arabic script, however, abbreviations do not use diacritic signs in both dotted and un-dotted styles [16]. For example, the abbreviations used for trigonometric functions contain at least two connected letters (see Table 5). Also, Arabic mathematical constants such as the canonical symbol for the radius of a circle {نق}, called Noq, and units abbreviations such as the symbol for kilogram measure (km) {كم or كلم} are written with cursive letters.

	English	Arabic
Abbreviation	sin	جا
	cos	جتا
	tan	ظا
Units	km	كم
	cm	سم
	mm	مم

Table (5): Example of Arabic abbreviations written cursively.

2.2.5. Numeral Systems in Arabic Mathematical Notation

In Arabic countries there are three numeral systems used, based on local preferences (see Table 6). These are:

- Arabic (European)
- Eastern Arabic (Arabic–Indic)
- Eastern Arabic-Indic

The European numeral system is also called the Western Arabic numeral system, the Arabic numeral system, or Maghreb Arab numbers. This system contains ten digits (0,1,2,3,4,5,6,7,8,9). These digits originated from the Hindu-Arabic numeral system, which was developed by Indian mathematicians. They were then adopted by Persian mathematicians in India, and delivered to Arab countries in the West.

There is some ambiguity regarding the naming of the European numeral system. Europeans refer to the system as the ‘Arab’ system (i.e. Arab numerals), since it was introduced to Europe by the Arabs of North Africa. On the other hand, in Arab countries the system is often referred to as ‘Indic’ or ‘Hindu’ (in addition to ‘Arabic’) due to its origin in India. [17].

The European numeral system is widely used by Arabs in North Africa, from Libya to Morocco. The Arabs of North Africa introduced the digits to Europe in the 10th century (see

Figure 17). From there, use of Arabic numerals spread around the world due to European trade, and the rise of printed publications [17].

The Eastern Arabic numeral system is also referred to as Arabic-Indic numerals, or Arabic Eastern numerals. It is used in the Middle East and Arabic countries in the east of Africa, such as Egypt [2].

Eastern Arabic-Indic numerals are used in Persia and in the Urdu mathematical language. The representation of these numbers is different from Eastern Arabic numerals (see Figure 17). For example, the numbers four and five are represented by different glyphs in Eastern Arabic-Indic than those used in Eastern Arabic.

System	Digit
1-Arabic numerals-European	0 1 2 3 4 5 6 7 8 9
2-Arabic-Indic	٩ ٨ ٧ ٦ ٥ ٤ ٣ ٢ ١ ٠
3-Eastern Arabic-Indic	۰ ۱ ۲ ۳ ۴ ۵ ۶ ۷ ۸ ۹

Table (6): The Numeral Systems Used in Arabic Mathematical Notation

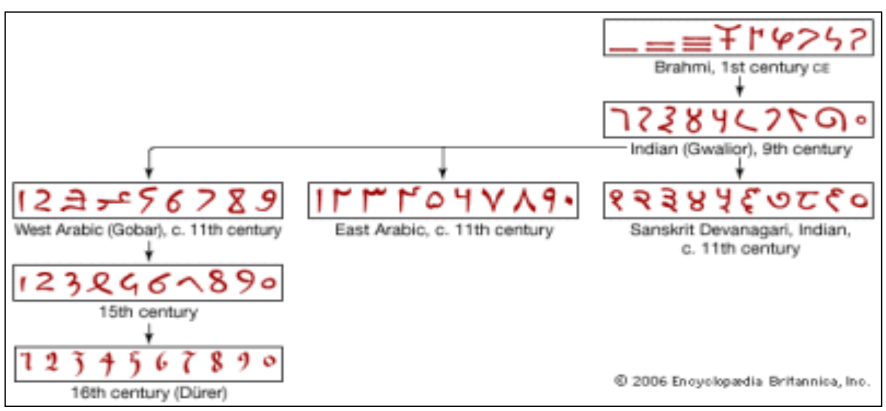


Figure (17): The Evolution of Hindu-Arabic Numerals
Source: http://www.storyofmathematics.com/indian.html

2.2.6. Direction of Arabic Expressions

Overall, Arabic writing flows from right to left. But, the Arabic language is called 'bi-directional' because while letters are written from right to left, numbers (possibly occurring in between letters) are written from left to right (see Figure 4).

2.2.7. Superscript and Subscript

In mathematical language, superscripts and subscripts are symbols are used to identify variables or numbers. Subscripts are used to label and indicate bases, variables, and indexes, and to eliminate ambiguity between numbers and variables in expressions (see Table 7) [18]. In the Arabic language, subscripts are represented either by a small symbol (such as a letter or number) written below the variable, or by a number on the left side of the variable.

Superscripts are used in various contexts in mathematics. For instance, superscripts are used to show the exponent or power when written above a number or variable, and if written between a function letter and its variable, the superscript represents the number of derivatives or iteration of the function [19]. The meaning of the superscript is determined by its context within the expression (see Table 8).

Indicate the Index	$\text{ب}_{\text{ن}} = \text{ب}_{\text{ن}-1} + \text{ب}_{\text{ن}-2}$ $\text{ب}_1 = 200$ ، $\text{ب}_2 = 300$ ، $\text{ب}_3 = 400$ ....
Label Variable	أ_1 , أ_2 ، ب_1 ، ب_2 ، جـ_1 ، جـ_2
Label Elements in Matrix	$\begin{pmatrix} \text{أ}_{٣١} & \text{أ}_{٢١} & \text{أ}_{١١} \\ \text{أ}_{٣٢} & \text{أ}_{٢٢} & \text{أ}_{١٢} \\ \text{أ}_{٣٣} & \text{أ}_{٢٣} & \text{أ}_{١٣} \end{pmatrix}_{٣\times٣} = \text{أ}_{٣\times٣}$

Table (7): Subscript Examples in Arabic Notation.
Source: http://www.schoolarabia.net/math/general_math/level3

Notation for Iterated Functions	$ق^{3}$ (س)
Notation for Exponent	$س^{4}$=س x س x س x س $س^{ن+م}$

Table (8): Superscript Examples in Arabic Notation.

2.2.8. Arabic Punctuation

Punctuation is a system of marks used together with letters and glyphs to supplement a writing system. Punctuation aids in making sentences more readable and understandable by eliminating ambiguity between phrases. This is in contrast to diacritic marks, which are used to represent the sound and pronunciation of letters [20].

The rules of punctuation systems vary from one language to another. For mathematical notation, the punctuation used in Arabic is different from that which is used in English (it is mirrored in Arabic since the writing direction is reversed).

There are four types of punctuation marks used in writing:

- **Stoppage Marks** such as (، ؛ .): these marks enable the reader to pause for a short time, and then continue reading.
- **Vocal Marks** such as (: ... ؟ !): these marks represent both stoppage and emotional tones to be used while reading.
- **Enclosing Marks** such as (« » - ( ) [ ]): these marks are used to organize words in order to make written text comprehensible.
- **Programming and Mathematical Marks** such as (< > * & ^ \ []).

Table 9 gives examples of the different uses of mathematical punctuation in Arabic and English. In English, the comma symbol (,) represents the limit between thousands, and the period symbol (.) is used to show fractions in decimal numbers. In the Arabic language, they are used the other way around.

	Arabic	English
Natural Numbers in Math	ط {…… ،3،2،1،0} {……،٣،٢،١,٠} Using ، to separate between the numbers	ℕ {0,1,2,3,…} Using **,** to separate between the numbers
Pi	٣,١٤١٥٩ 3,14159 Using , in the decimal Numbers	3.14159 Using point or dot in decimal numbers
Representation for Numbers of Four or more Digits	١.٠٠٠ 1.000	1,000

Table (9): Examples showing differences between Arabic and English mathematical punctuation marks.

<u>2.2.9. Mirrored Symbols</u>

Symbols that are used in most languages, such as parentheses and quotation marks, can be handled uniformly by typesetting systems. In Arabic, some of these symbols are simply mirror images of their English form. So, to avoid redundancy, such symbols are omitted from the Unicode blocks for Arabic. For example, in Arabic there are symbols with the same shape as in English, such as ‘Σ‘, but which are mirrored in Arabic writing, as in  [2]. The corresponding Unicode character, when used, is marked as ‘mirrored’. Table 10 gives examples of mathematical operation symbols, and their corresponding mirrored symbols.

Symbol Name	Character Code	English	Arabic
THERE EXISTS	U+2203	∃	E
N-ARY SUMMATION	U+2211	Σ	Ƹ
ELEMENT OF	U+2208	∈	∋
INTEGRAL	U+222B	∫	ʅ
NOT A SUBSET OF	U+2284	⊄	⊅
NOT A SUPERSET OF	U+2285	⊅	⊄
Angle	U+2220	∠	⦣

Table (10): Examples of mathematical operator symbols used in English and Arabic.

2.2.10. Ligature Shaping (Calligraphic Symbols)

The Arabic mathematical language uses ligature symbols. These are operators created from connecting two or more sequential letters. For instance, the limit operator is represented by a ligature comprising three connected letters (see Tables 11 & 12). The layout of ligature symbols is different from regular text containing the same letters. Ligatures have curvilinear kashida and their size is adjusted according to the overall expression (see Figure 18) [16].

Isolated Letter	ا ALEF	ه HAH	ن NON
Contextual Form (Letter Position)	ـا FINAL	ـهـ MEDIAL	نـ INITIAL
Ligature Symbol	نها		

Table (11): The Limit Operator Ligature in Arabic Mathematical Language

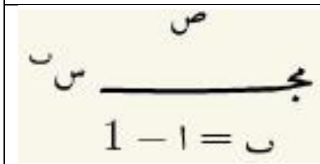
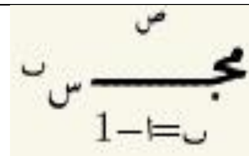
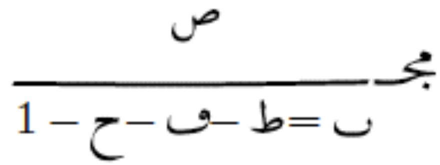
Curvilinear Kashida	Linear Kashida	Variable Size Kashida
		

Figure (18): Summation operator has variable size
Source: http://www.ucam.ac.ma/fssm/rydarab/

Symbol Name	English	Arabic
Limit	lim	نهـــا حـــد
Product	$\prod$	جـــذ $\prod$ in some regions
Summation	Σ	مجـــــــ Ƹ in some regions
Logarithm	log	لو
Factorial	!	! in some regions ل لـ

Table (12): Examples showing mathematical ligature symbol operators used in Arabic and corresponding English symbols.

3. Mathematical Notation in Arabic Regions

Within Western Latin languages, mathematical notation evolved over time as knowledge grew, and in particular with the advent of modern mathematics. Along the way, typesetting systems expanded accordingly and aided in keeping the Latin notation uniform. Arabic mathematical notation had not changed for centuries. By the time classrooms were ready for modern mathematics, foreign languages such as Arabic required extended mathematics notation. Regional variations arose as modern concepts were adopted into educational curricula. The result is that Arabic mathematics notation is partially, if not completely, localized to specific Arab countries [14].

Mathematical notation in the Arabic language changes from one Arab country to another. Mathematics books published in Morocco are not used in Egypt (and vice versa) because the direction of writing formulas and symbols is different. These differences are reflected in the way word processing programs distinguish between Arab Jordan, Syria, Morocco and Egypt.

According to cultural influences in each region of the Arab countries, Arabic notation can be classified as belonging to one of two styles:

1. Dual mathematical presentation (where mathematics is presented in English or French), or
2. Arabic mathematical presentation

In the former style, the symbols used are taken from one European language (English or French); the choice is determined by the level of cultural influence of each language on the region. For instance, in the second row of Table 12, notation is written from left to write and the symbols used are taken from French notation. This is an example of the notation used in Morocco, and it is called Moroccan style (which is influenced by French math notation). Formulas are written from right to left. Symbols are written in the opposite direction, as with text in the host natural language (Arabic), and expressions are intermixed and bidirectional (see Table 13) [14].

Influences Culture	Influence Culture Formula	Arabic Country Formula	
English	$f(x) = \begin{cases} \sum_{i=1}^{s} x^i & \text{if } x < 0 \\ \int_1^s x^i \, dx & \text{if } x \in S \\ \tan \pi & \text{otherwise (with } \pi \simeq 3.141) \end{cases}$	$f(x) =$ { $\sum_{i=1}^{s} x^i$ إذا كان $x < 0$ $\int_1^s x^i \, dx$ إذا كان $x \in S$ $\tan \pi$ غير ذلك(مع $\pi \simeq 3.141$)	Saudi Arabia, Kuwait
French	$f(x) = \begin{cases} \sum_{i=1}^{s} x^i & \text{si } x < 0 \\ \int_1^s x^i \, dx & \text{si } x \in E \\ \text{tg}\, \pi & \text{sinon (avec } \pi \simeq 3{,}141) \end{cases}$	$f(x) =$ { $\sum_{i=1}^{s} x^i$ إذا كان $x < 0$ $\int_1^s x^i \, dx$ إذا كان $x \in E$ $\text{tg}\ \pi$ غير ذلك ($\pi \simeq 3,141$ مع)	Moroccan style (which is influenced by French math notation)

Table (13): Dual Mathematical Presentation

In the second style, notation is written from right to left, which is the writing direction of the natural Arabic language. For numbers, this style uses either Arabic European or Arabic-Indic digits. Operator symbols can be either specific symbols from the Arabic alphabet or mirrored symbols which are inherited from Latin languages. For instance, the summation operator (sigma) is represented by

مجـــــ or ∑ (mirrored from the Latin summation operator symbol) (see Figure 19) [14].

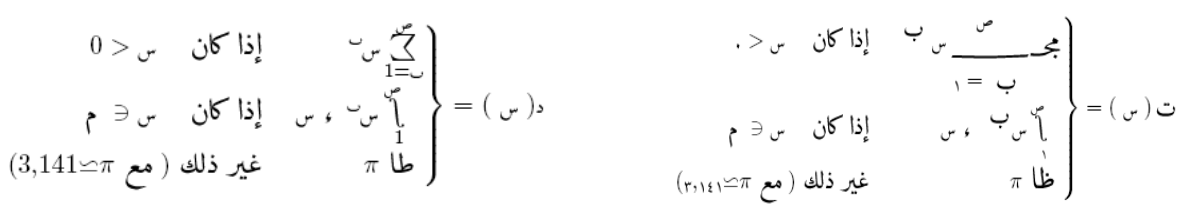


Figure (19): Arabic Mathematical Presentation

Arabic mathematical presentation is further classified into two types: Mashrek and Maghreb. The Mashrek style of mathematical notation is written from right to left, uses Arabic-Indic numerals, and uses Arabic letters for some operator literal symbols (such as the summation, limit, product, and logarithm operators). Other symbols such as ELEMENT OF and INTEGRAL are written in the direction of Arabic writing and mirrored. These operators are available in Unicode and are marked as mirrored. The Mashrek style is commonly used in Middle Eastern countries.

The Maghreb style of mathematical notation is written from right to left, uses both Arabic-Indic or European numerals, and uses mirrored Latin notation for symbols such as ELEMENT OF and INTEGRAL.

4. Discussion

Individual efforts to incorporate local Arabic mathematical notations into software have succeeded. But, language standards are required in order for the widely used word processing programs to consistently add support for localized Arabic notation. Only then can Arab authors have access to the same knowledge and standards for technical communication as are available in English, allowing them to publish scientific works and join the global pool of scientific references.

Standards development organizations have collaborated in seeking to represent and define rules for handling the characters of all international languages. The resulting Unicode standard contains several metrics for Arabic calligraphy, but for which consensus has failed to be reached on some elements. Moreover, the Unicode Consortium has yet to add domain-specific Arabic characters such as those used in mathematics [21].

For Arab science to improve, the Arabic language must be adopted for teaching science at all levels, from early schooling to higher education [22]. The prerequisites for enabling the writing of scientific text in localized Arabic notation include:

1. The production of Arabic mathematical fonts, including all symbols used in Arabic countries, with homogeneous shapes, sizes, and boldness.
2. The addition of Arabic mathematical symbols to the Unicode encoding standard.
3. The adoption by MathML of international standards for structuring Arabic formulas.

The purpose of adopting Arabic formula structure within the W3C's MathML specification is to encourage an international standard for Arabic formula writing. MathML is an XML markup language that is used to build mathematical expressions that can be displayed, manipulated, and shared over the World Wide Web. Expressions encoded in MathML can be presented on any device with a Web browser, cut and pasted into a word processor, and printed on a laser printer [23]. We now discuss MathML in more detail.

4.1. The MathML Software Tool

MathML is an application of XML for presenting mathematical notation as well as representing its structure and content. MathML is ideal for use by any software application that involves mathematical expressions. There are a number of implementations of MathML as well as software that supports and uses it [4] (see Table 14).

Category	Example	Vendor	
Browser	Dadzilla	RyDArab	Adapted version of Mozilla, Using MathML for Arabic mathematical presentation
Browser Plug in	MathPlayer	Design Science	MathML engine for Microsoft's Internet Explorer browser to fast track the adoption of MathML in science
Editor	WIRIS	WIRIS	Based on MathML and Support LaTeX and Arabic Writing from right to left
Scientific Computation	Maple	Maple Soft	Scientific Computation System support MathML
Composition and Rendering Engine	MathMagic	InfoLogic, Inc.	Convert from/to MathML and TeX.
Converter	Translator from Tex to MathML	ORCCA	Convert mathematical expressions from TeX/LaTeX to MathML
Authoring System	GELLMU		Generate accessible XHTML+MathML
Stylesheet to/from MathML	DSSSL		Parse the content and presentation of the MathML and render it to TeX or rtf.
DTDs Schema	XML Schema for MathML	W3C	Modular XML Schema for MathML
Component and SDKs	MathML.Net	Soft4science	MathML component implemented on .Net Framework and provide integration on Visual Studio.Net
Research Project	Enhancing Searching of Mathematics Project	Design Science	Improved search for documents that containing mathematics
Accessibility	LAMBDA	LAMBDA TEAM	Access to Mathematic for Braille Device

Table (14): Examples of software that use or implement MathML

4.2. Examples of Arabic Mathematical Expressions Written in MathML

In this section, we give several examples of Arabic mathematical expressions and their corresponding MathML representations. The MathML was displayed in Mozilla Firefox Explorer (29.0.1) on the Mac OS X version 10.7.5.

	Image	MathML
1	$f(x) = \begin{cases} \sum_{i=1}^{s} x^i & \text{if } x<0 \\ \int_1^s x^i\,dx & \text{if } x\in S \\ \tan\pi & \text{otherwise (with } \pi\simeq 3.141) \end{cases}$ **English Style**	<math xmlns="http://www.w3.org/1998/Math/MathML" display="block"> <mrow> <mrow> <mi>f</mi> <mrow> <mo>(</mo> <mi>x</mi> <mo>)</mo> </mrow> </mrow> <mo>=</mo> <mrow> <mo>{</mo> <mtable> <mtr> <mtd> <mrow> <munderover> <mo movablelimits="false">∑</mo> <mrow> <mi>i</mi> <mo>=</mo> <mn>1</mn> </mrow> <mi>s</mi> </munderover> <msup> <mi>x</mi> <mi>i</mi> </msup> </mrow> </mtd> <mtd> <mrow> <mtext> if </mtext> <mi>x</mi> <mo><</mo> <mn>0</mn> </mrow> </mtd> </mtr> <mtr> <mtd> <mrow> <msubsup> <mo>∫</mo> <mn>1</mn> <mi>s</mi> </msubsup> <mrow> <msup> <mi>x</mi> <mi>i</mi> </msup> <mi>d</mi>

```
        <mi>x</mi>
       </mrow>
      </mrow>
     </mtd>
     <mtd>
      <mrow>
       <mtext> if </mtext>
       <mi>x</mi>
       <mo>∈</mo>
       <mi mathvariant="normal">S</mi>
      </mrow>
     </mtd>
    </mtr>
    <mtr>
     <mtd>
      <mrow>
       <mi>tan</mi>
       <mi>π</mi>
      </mrow>
     </mtd>
     <mtd>
      <mrow>
       <mtext> otherwise </mtext>
       <mrow>
        <mo>(</mo>
        <mtext>with </mtext>
        <mi>π</mi>
        <mo>≃</mo>
        <mn>3.141</mn>
        <mo>)</mo>
       </mrow>
      </mrow>
     </mtd>
    </mtr>
   </mtable>
  </mrow>
 </mrow>
</math>
```

2

$$f(x) = \begin{cases} \sum_{i=1}^{s} x^i & \text{si } x < 0 \\ \int_1^s x^i \, dx & \text{si } x \in \mathrm{E} \\ \mathrm{tg}\,\pi & \text{sinon (avec } \pi \simeq 3{,}141) \end{cases}$$

French Style

```
<math xmlns="http://www.w3.org/1998/Math/MathML" display="block">
 <mrow>
  <mrow>
   <mi>f</mi>
   <mrow>
    <mo>(</mo>
    <mi>x</mi>
    <mo>)</mo>
   </mrow>
  </mrow>
  <mo>=</mo>
  <mrow>
   <mo>{</mo>
   <mtable>
    <mtr>
     <mtd>
      <mrow>
       <munderover>
```

```
        <mo movablelimits="false">∑</mo>
        <mrow>
          <mi>i</mi>
          <mo>=</mo>
          <mn>1</mn>
        </mrow>
        <mi>s</mi>
      </munderover>
      <msup>
        <mi>x</mi>
        <mi>i</mi>
      </msup>
    </mrow>
  </mtd>
  <mtd>
    <mrow>
      <mtext> si  </mtext>
      <mi>x</mi>
      <mo>&lt;</mo>
      <mn>0</mn>
    </mrow>
  </mtd>
</mtr>
<mtr>
  <mtd>
    <mrow>
      <msubsup>
        <mo>∫</mo>
        <mn>1</mn>
        <mi>s</mi>
      </msubsup>
      <mrow>
        <msup>
          <mi>x</mi>
          <mi>i</mi>
        </msup>
        <mi>d</mi>
        <mi>x</mi>
      </mrow>
    </mrow>
  </mtd>
  <mtd>
    <mrow>
      <mtext> si </mtext>
      <mi>x</mi>
      <mo>∈</mo>
      <mi mathvariant="normal">E</mi>
    </mrow>
  </mtd>
</mtr>
<mtr>
  <mtd>
    <mrow>
      <mi>tg</mi>
      <mi>π</mi>
    </mrow>
```

		</mtd> <mtd> <mrow> <mtext> sinon </mtext> <mrow> <mo>(</mo> <mtext>avec </mtext> <mi>π</mi> <mo>≃</mo> <mn>3,141</mn> <mo>)</mo> </mrow> </mrow> </mtd> </mtr> </mtable> </mrow> </mrow> </math>
3	$\text{د(س)} = \begin{cases} \sum_{\text{ب}=1}^{\text{ص}} \text{س}^{\text{ب}} & \text{اذا كان س} > 0 \\ \int_{1}^{\text{ص}} \text{س}^{\text{ب}} \text{دس} & \text{اذا كان س} \ni \text{م} \\ \text{طا}\pi & \text{غير ذلك(مع}\ \pi \simeq 3,14) \end{cases}$ Arabic Maghreb Style	<math dir ="rtl" xmlns="http://www.w3.org/1998/Math/MathML" display="block" > <mrow> <mrow> <mi>د</mi> <mrow> <mo>(</mo> <mi>س</mi> <mo>)</mo> </mrow> </mrow> <mo>=</mo> <mrow> <mo>{</mo> <mtable> <mtr> <mtd> <mrow> <munderover> <mo movablelimits="false">∑</mo> <mrow> <mi>ب</mi> <mo>=</mo> <mn>1</mn> </mrow> <mi>ص</mi> </munderover> <msup> <mi>س</mi> <mi>ب</mi> </msup> </mrow> </mtd> <mtd> <mrow> <mtext> اذا كان </mtext> <mi>س</mi>

		<mo><</mo> <mn>0</mn> </mrow> </mtd> </mtr> <mtr> <mtd> <mrow> <msubsup> <mo>∫</mo> <mn>1</mn> <mi>ص</mi> </msubsup> <mrow> <msup> <mi>س</mi> <mi>ب</mi> </msup> <mi>د</mi> <mi>س</mi> </mrow> </mrow> </mtd> <mtd> <mrow> <mtext> اذا كان </mtext> <mi>س</mi> <mo>∈</mo> <mi mathvariant="normal">م</mi> </mrow> </mtd> </mtr> <mtr> <mtd> <mrow> <mi> 𞹨 </mi> <mi>π</mi> </mrow> </mtd> <mtd> <mrow> <mtext> غير ذلك </mtext> <mrow> <mo>(</mo> <mtext><rl>مع </rl> </mtext> <mi>π</mi> <mo>≃</mo> <mn>3,141</mn> <mo>)</mo> </mrow> </mrow> </mtd> </mtr> </mtable> </mrow> </mrow>

<table>
<tr><td></td><td></td><td></math></td></tr>
<tr><td>4</td><td>ت (س) = {
مجـــــ ص ب = ١ س ب إذا كان س < ٠
∫ ١ ص س ب ، س إذا كان س ∈ م
ظا π غير ذلك (مع π≃٣٫١٤١)
Arabic Mashrek Style</td><td><math dir ="rtl" xmlns="http://www.w3.org/1998/Math/MathML" display="block" >
<mrow>
<mrow>
<mi>ت</mi>
<mrow>
<mo>(</mo>
<mi>س</mi>
<mo>)</mo>
</mrow>
</mrow>
<mo>=</mo>
<mrow>
<mo>{</mo>
<mtable>
<mtr>
<mtd>
<mrow>
<munderover>
<mo>𞻰</mo>
<mrow>
<mi>ب</mi>
<mo>=</mo>
<mn>١</mn>
</mrow>
<mi>ص</mi>
</munderover>
<msup>
<mi>>mi/<س
<mi>ب</mi>
</msup>
</mrow>
</mtd>
<mtd>
<mrow>
<mtext> اذا كان </mtext>
<mi>س</mi>
<mo><</mo>
<mn>٠</mn>
</mrow>
</mtd>
</mtr>
<mtr>
<mtd>
<mrow>
<msubsup>
<mo>∫</mo>
<mn>1</mn>
<mi>ص</mi>
</msubsup>
<mrow></td></tr>
</table>

		<msup> <mi>س</mi> <mi>ب</mi> </msup> <mi>ء</mi> <mi>س</mi> </mrow> </mrow> </mtd> <mtd> <mrow> <mtext> اذا كان </mtext> <mi>س</mi> <mo>∈</mo> <mi mathvariant="normal">م</mi> </mrow> </mtd> </mtr> <mtr> <mtd> <mrow> <mi> 𞹺 </mi> <mi>π</mi> </mrow> </mtd> <mtd> <mrow> <mtext> غير ذلك </mtext> <mrow> <mo>(</mo> <mtext><rl>مع </rl> </mtext> <mi>π</mi> <mo> ≃ </mo> <mn>٣,١٤</mn> <mo>)</mo> </mrow> </mrow> </mtd> </mtr> </mtable> </mrow> </mrow> </math>

Table (15): Representing Arabic Notation in MathML

Example 3 in Table 15 depicts the use of mirrored symbols within the Arabic Maghreb style. Mirrored symbols are required for compatibility with the right-to-left writing direction. Symbols are mirrored by marking instances of their Unicode encoding as mirrored. Within the expression, the Arabic symbol DAL (د) is the initial letter of دالة or "function" in Arabic. The Arabic letter BEH ٮ (dot-less form of the Arabic letter ب) and the letters of the function name abbreviation, طا , are written without dots. Note also that the number format can use either the Arabic-Indic or European numeral systems [2].

It can be seen in Example 4 of Table 15 that the Mashrek style of writing for the expression differs from the Maghreb style of Example 3. The Arabic symbol مجـ is used to denote summation (the initial character of مجموع , i.e. "sum" in Arabic). The letter ت is used for the function (the initial character of تابعة, i.e. "function" in Arabic). Further, the letters of the function name abbreviation (ظا) are written with dots, and the number format uses Arabic-Indic digits. The comma for the decimal separator is the same in both styles [2].

4.3 Arabic Mathematical Localization Implementing Issues

The first two versions of MathML (1.0 and 2.0) were designed to represent and describe mathematics in European languages. Subsequently the W3C has released MathML 3.0, which includes international mathematical notations. In this section, we aim to illustrate how the enhancements embodied in MathML 3.0 are capable of describing Arabic mathematical notation

4.3.1. Illustration of Expression with Arabic Writing Direction

The Unicode Bidirectional Algorithm (Bidi) is used by browsers in order to display text in the correct order. The algorithm provides a means for handling a document's text direction without requiring that special markup or other mechanisms be added to the document. However, the algorithm does not address documents containing a mix of both text directions. To handle bidirectional text documents, additional markup is required [24].

In MathML, expressions are displayed from left to right and bi-directionality is not supported in text. For example, consider the expression س + ٢- ١, which is Arabic notation for 1-x+2. When the Arabic notation is marked up in MathML in the Arabic writing direction (i.e. right-to-left order), the expression is rendered backwards in the browser (see Table 16). The expression is incorrectly displayed as ١ – س+ ٢ (which when read from RTL is equivalent to 2+x-1 in English).

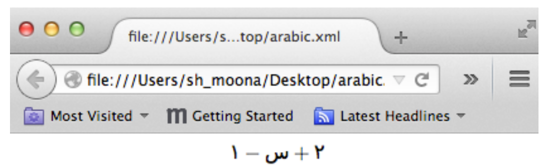
MathML Code	Output
<math xmlns='http://www.w3.org/1998/Math/MathML' display='block'> <mrow> <mn> ١ </mn> <mo> - </mo> <mi> س </mi> <mo> + </mo> <mn> ٢ </mn> </mrow> </math>	

Table (16): Representing Arabic Notation from RTL.

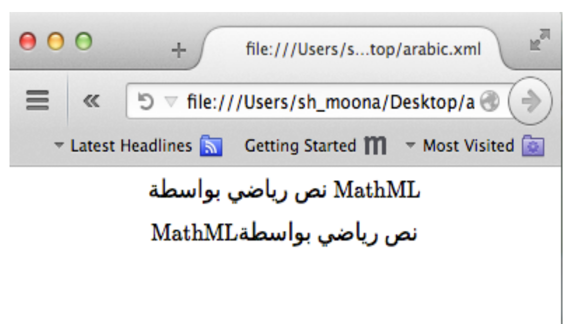
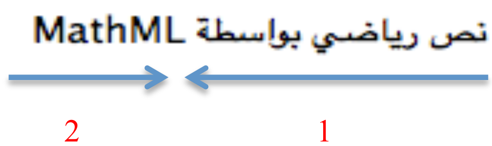
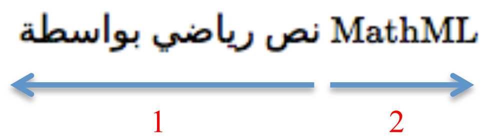
MathML Code	Output
<math xmlns='http://www.w3.org/1998/Math/MathML' display='block'> <mtable> <mtr> <mrow> <mtext>نص رياضي بواسطة MathML </mtext> </mrow> </mtr> <mtr> <mrow> <mtext> MathML نص رياضي بواسطة </mtext> </mrow> </mtr> </mtable> </math>	

Table (17): Representing bidirectional notation written in MathML using <mtext>.

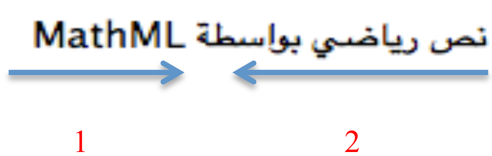
In Table 17, the first <mtext> element implicitly contains both writing directions (RTL followed by LTR). The writing begins from the right side for the Arabic letters and then switches to the left side for the English letters. The desired display text (reading is intended to begin from RTL) is:



However, the MathML is rendered as:



The second <mtext> element implicitly contains both writing directions as well, this time beginning with LTR for the English letters and then switching to RTL for the Arabic letters. The desired display text (reading is intended to begin from LTR) is:



The MathML for this mixed-direction sequence is rendered the same as for the first sequence.

For MathML elements that can contain text (<mtext>, <mo>, <mi>, <mn>, and <ms>), the direction of rendering for the text content is determined by its Unicode properties. If the text contains mixed writing directions (as with the above examples), the Bidi algorithm must be used [24]. Tables 18 and 19 give examples of how the Bidi algorithm works in different cases of mixed-direction character sequences.

We can see, from Tables 17, 18, and 19 that the overall writing direction of the text inside each element is assumed to be LTR. That is, the base direction is LTR, which is inherited from the default direction of the document. So, applying the Bidi algorithm to the text in each element works only for text that is to be read from LTR. An attribute needs to be added to these elements in order to be able to change the default direction of each element or the overall document.

Logical Order	Visual Order
ab	ab
ب ا	ب ا
a ب ا c d	a ا ب c d
ج د a ب	د ج a ب

Table (18): Representing changes in writing direction assuming the overall document direction is LTR

Logical Order	Visual Order
ab	ab
ب ا	ب ا
a ب ا c d	c d ا ب a
ج د a ب	ب a ج د

Table (19): Representing changes in writing direction assuming the overall document direction is RTL

In Tables 18 and 19, the logical order is the order in which the letters are keyed and stored in memory, while the visual order is the order in which the letters are presented on the screen or page [24].

MathML uses the XML standard for expressing mathematical notations. In MathML 3.0, the W3C added support for Arabic mathematical notation in Unicode by adding a direction attribute. The “dir” attribute is used to indicate whether the writing direction of the text within an element is LTR or RTL. Firefox (which uses Mozilla's layout engine) and Dadzilla, the two browsers that support Arabic notation, subsequently implemented the RTL feature [26]. The Dadzilla browser is an improved version of Mozilla, which supports MathML for Arabic mathematical presentation [4].

In MathML 3.0, there are two different directions that can be used in order to represent mathematical notation:

1. The overall layout schemata (i.e. the direction of the mathematical formulas)
2. The direction of the text and symbols contained in elements

The default direction of MathML documents is LTR. The layout schemata direction (i.e. overall direction of the document) is determined by the ‘dir’ attribute inside the <math> element, which is the top-level element in MathML expressions (see Table 20). For instance, if “dir” is set to "rtl", then the overall layout is from RTL. Subscripts and superscripts appear to the left of the base, and mirror symbols from the Unicode block for the conventional European layout are used. However, dual mathematical presentation can contain individual subformulas with directions different from their container element. For such subformulas, the “dir” attribute can be used on <mrow> or <mstyle> elements to change the direction set in the <math> element [24].

Image	MathML
<math xmlns='http://www.w3.org/1998/Math/MathML' display='block' dir='rtl'> <mtext> نص رياضي باللغة العربيه </mtext> <mrow> <mo> ∑ </mo> <mi> س </mi> <mn> ٥ </mn> </mrow> </math>	نص رياضي باللغة العربيه ∑ س٥
Arabic Style from RTL	
<math xmlns='http://www.w3.org/1998/Math/MathML' display='block' dir='rtl'> <mtext> نص رياضي باللغه العربيه بواسطة MathML </mtext> <mrow> <mo> ∑ </mo> <mi> س </mi> <mn> ٥ </mn> </mrow> </math>	نص رياضي باللغه العربيه بواسطةMathML ∑ س٥
Arabic Style from RTL with English Text	
<math xmlns='http://www.w3.org/1998/Math/MathML' display='block'> <mtext> MathML </mtext> <mrow> <mi> sin </mi> <mn>45</mn> </mrow> </math>	MathML sin45
<math xmlns='http://www.w3.org/1998/Math/MathML' display='block' dir='rtl'> <mtext> نص رياضي MathML</mtext> <mrow> <mi> جتا </mi> <mn>٤٥</mn> </mrow> </math>	نص رياضيMathML جتا٤٥

Table (20): Examples of using dir="rtl" in overall document direction.

<table>
<tr><th>Image</th><th>MathML</th></tr>
<tr><td><math
xmlns='http://www.w3.org/1998/Math/MathML'
display='block'>
<mtext dir='rtl'> نص رياضي MathML</mtext>
<mrow>
<mo> ∑ </mo>
<mn> 45 </mn>
</mrow>
</math></td><td>MathMLنص رياضي ∑45</td></tr>
<tr><td colspan="2">LTR layout with mixed text dir=rtl</td></tr>
<tr><td><math
xmlns='http://www.w3.org/1998/Math/MathML'
display='block' dir='rtl'>
<mtext dir='ltr'> نص رياضي
MathML</mtext>
<mrow>
<mo> ∑ </mo>
<mn> 45 </mn>
</mrow>
</math></td><td>45 ∑ نص رياضي MathML</td></tr>
<tr><td colspan="2">RTL layout with mixed text dir=ltr</td></tr>
<tr><td><math
xmlns='http://www.w3.org/1998/Math/MathML'
display='block' dir='rtl'>
<mtext> نص رياضي محتوي على دوال مثلثيه
>mtext/<جتا
<mi> جتا </mi>
</math></td><td>نص رياضي محتوي على دوال مثلثيه جتا جتا</td></tr>
<tr><td colspan="2">RTL layout with Arabic Trigonometric functions</td></tr>
<tr><td><math
xmlns='http://www.w3.org/1998/Math/MathML'
display='block' dir='rtl'>
<mtext> نص رياضي محتوي على دوال مثلثيه
>mtext/<جتا
<mi><rl> جتا </rl></mi>
</math></td><td>نص رياضي محتوي على دوال مثلثيه جتا جتا</td></tr>
<tr><td colspan="2">RTL layout with Arabic and used <rl>(see Table 20) element inside <mi></td></tr>
</table>

Table (21): Examples of using dir=rtl in <mi> and text tokens.

For changing the direction of text and symbols, MathML includes elements that can contain text (<mtext>, <mo>, <mi>, <mn>, and <ms>) and Unicode is used to represent this text. If the element text contains only LTR or RTL characters, the text is displayed in the given direction.

However, if mixture of directions is desired then the Unicode Bidi algorithm must be applied. The Bidi algorithm is run on the contents of each element (see Tables 18 and 19) [24]. Table 21 gives examples of the use of the “dir” attribute inside text-containing elements (compare with the examples given in Tables 18 and 19, in which the overall direction is LTR). Table 21 gives examples of formulas that contain general text. Text is used within mathematical expressions to comment formulas in whole or in part, and is represented in MathML using the <mtext> element.

The first row of Table 21 shows MathML code for an LTR math document containing bidirectional text. In this case, the Bidi algorithm can be used to determine directionality of the text since the overall document direction is LTR.

In the second row of Table 21, the overall document direction is RTL and the mixed text is presented from LTR.

In the third and fourth rows of Table 21, the layout of the document is RTL. It should be noted that the <mi> element is not useful for representing Arabic abbreviations for trigonometric functions, since these abbreviations are written as ligatures (i.e. two joined letters; see Table 12) and <mtext> and <ms> are the only elements that support cursive writing.

The <rl> element, a proposed extension to MathML, is similar to the <mrow> element, but renders text from RTL. We did use the <rl> element with the <mi> element to connect function-name letters and achieve the RTL arrangement of any sub-expression within an expression; however, the <rl> element cannot be used within the <msub> and <msup> elements, or any other element that requires more than one argument. The <rl> element counts as one argument when used within these elements. For this reason, two new elements, <amsub> and <amsup> need to be added to MathML markup. They are used within formulas written from RTL and have the same size as the <msub> and <msup> elements (see Table 22) [14].

As mentioned in [27], the element <rl> is not a useful solution for representing direction because the encoding of Arabic notation becomes heavier. The aim of using MathML is to represent an expression without affecting its meaning. Taking this into account, any new elements have to be visibly understandable for the user.

MathML Code	Output
<math xmlns="http://www.w3.org/1998/Math/MathML" display="block" > <msup> <rl> <mi>س</mi> <mi>ب</mi> </rl> </msup> </math>	invalid-markup
<math xmlns="http://www.w3.org/1998/Math/MathML" display="block" > <amsup> <mi>س</mi> <mi>ب</mi> </amsup> </math>	س ب
<math xmlns="http://www.w3.org/1998/Math/MathML" display="block" dir='rtl'> <msup> <mi>س</mi> <mi>ب</mi> </msup> </math>	$س^{ب}$

Table (22): Using <amsup> and <msup> with RTL text

In Table 22, we can see rendering results corresponding to an implementation of the <amsup> element, which is added to MathML markup in order to represent Arabic superscript notation (replacing <msup>). Here we can see that <amsup> does not represent Arabic <msup> expressions correctly. Using the ‘dir=rtl’ attribute value, however, results in the correct presentation.

In Arabic mathematical notation some operator symbols, such as square root and summation (see Table 10), are simply mirror images of the corresponding Latin ones. Thus, a new font family is required to represent them. The Arabic Computer Modern (ACM) fonts have been introduced to provide Computer Modern fonts for Arabic that are in the TrueType format and have mirror glyphs symbols. In [12] a new MathML element, <amath>, is proposed for using the ACM fonts. The <amath> element is not required if the Arabic mathematical symbols are already contained in the Unicode tables and marked as mirrored. Moreover, the direction of text contained in an <amath> element should be RTL, since ACM fonts are used to present it. But, in MathML the default direction of the document is LTR, which cannot be changed by adding ‘dir=rtl’ to the <amath> element. So, to change the direction of presenting a formula within <amath>, the <rl> element would be required and it would need to be used for each sub-expression (see Table 23).

<table>
<tr><th>MathML Code</th><th>Output</th></tr>
<tr><td rowspan="2"><amath
xmlns="http://www.w3.org/1998/Math/MathML"
display="block" >
<mi>س</mi>
<mo>+</mo>
<mi>ب</mi>
<mtext> معادلة رياضيه</mtext>
</amath></td><td>معادلة رياضيه ب + س</td></tr>
<tr><td>The layout direction is LTR</td></tr>
<tr><td rowspan="2"><math
xmlns="http://www.w3.org/1998/Math/MathML"
display="block" >
<mi>س</mi>
<mo>+</mo>
<mi>ب</mi>
<mtext> معادلة رياضيه</mtext>
</math></td><td>معادلة رياضيه ب + س</td></tr>
<tr><td>The layout direction is LTR</td></tr>
</table>

<amath xmlns="http://www.w3.org/1998/Math/MathML" > <mrow> <rl> <mi>س</mi> <mo>+</mo> <mi>ب</mi> </rl> </mrow> </amath>	ب + س The layout direction is LTR
<amath xmlns="http://www.w3.org/1998/Math/MathML" > <mrow> <rl> <mi>س</mi> <mo>+</mo> <mi>ب</mi> </rl> </mrow> <mtext><rl>نص رياضي عربي mathml</rl></mtext> </amath>	mathml نص رياضي عربي ب + س The layout direction is LTR
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <mrow> <mi>س</mi> <mo>+</mo> <mi>ب</mi> </mrow> <mtext> نص رياضي عربي mathml</mtext> </math>	س + ب نص رياضي عربيmathml The layout direction is RTL
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <amath> <mrow> <mi>س</mi> <mo>+</mo> <mi>ب</mi> </mrow> <mtext> نص رياضي عربي mathml</mtext> </amath> </math>	س + ب نص رياضي عربيmathml

Table (23): Using <math> and <amath>

4.3.2. Arabic Mathematical Symbols in Unicode and MathML

In order to support the publication of mathematics on different systems, the Unicode standard must define the symbols that are used in mathematical notation. Some Arabic mathematical symbols, such as the summation operator (see Table 10), are not defined separately in Unicode because they are mirror images of the mathematical symbols used in Latin script, which are already defined.

Mirrored symbols are not the solution in all cases. In a project entitled "Arabic Typesetting in the Electronic Scientific Multilingual", researchers at the Cadi Ayyad University in Morocco proposed the addition to Unicode of a set of mirror-image Arabic symbols, which were eventually taken up by the standard [28].

The mirrored punctuation symbol corresponding to the Latin comma (", ") is used in Arabic notation to separate thousands in numbers. The mirrored form is "،", for which no Latin symbol in the Unicode Bidi table is marked as mirrored (see Table 24). Thus, the only way to use this Arabic mathematical punctuation symbol is via a keyboard for the Arabic language.

The Latin mathematical symbol for summation (code point U+2211) is not distinct from the corresponding Arabic symbol. When used in Arabic expressions, the symbol is marked as mirrored (and shares the same Unicode code point). Meanwhile, in the Greek language, the letter Sigma is exactly the same as the Latin mathematical symbol for summation and yet it is treated as distinct (code point U+03A3). The Latin capital letter Esh is also represented by the same symbol, but is similarly assigned a unique code point of U+01A9 [2].

Mirrored Unicode counterparts of Latin symbols do not correctly represent their corresponding Arabic mathematical operations. For example, the European symbol $\not\subset$ is used for 'not a subset of' and has code point U+2284. Its mirror image, ⊅ at code point U+2285, could be used to represent 'not a subset of' in Arabic; however, this symbol actually represents the 'not a superset of' relation.

As of Unicode version 6.1, new characters are contained in the table for Arabic mathematical symbols. One of them, called the "stretch" operator, can be used to stretch expressions such as limits i.e. ـــــ (code point &#126705). Unfortunately, the Windows and Macintosh operating systems do not yet support such new code points, and so attempts to use them in browsers lead to undefined characters.

String literals (interpreted, for example, by programming languages and computer algebra systems) are demarcated in MathML with the <ms> element. When displayed, string literals are enclosed in double quotation marks [2]. The 'lquote' and 'rquote' attributes of the <ms> element can, however, be used to set custom display characters in place of the default double quotes.

The 'lspace' and 'rspace' attributes of <ms> are used to specify space before and after an operator (see Table 24). The unit used to measure space is called the 'EM', which is the font unit used for measuring horizontal lengths [2]. In order to size and align glyphs, the EM size is determined by an imaginary square. The boundary of the box completely surrounds all glyphs and contains the glyph ascent, descent, and extra spacing for fixing tangled lines (see Figure 20) [29].

MathML recommends that the lspace, rspace, lquote, and rquote attributes be rendered as opening and closing, to eliminate the ambiguity of using them with RTL notation. Thus, the standard Unicode names for parentheses, which are LEFT PARENTHESIS and RIGHT PARENTHESIS, are marked as mirrored. As mentioned above (for 'not a subset of'), the mirrored parentheses symbols do not accurately describe the intended RTL meaning of each.

From the data we collected from different mathematics handbooks, we found that there are Arabic symbols that are not included in Unicode and have not yet been proposed as additions due to the adoption of substitute symbols. For instance, the symbol for parallel lines in Unicode is ‖ , while the one used in Arabic notation is //.

Code	Output
<math xmlns="http://www.w3.org/1998/Math/MathML" dir="rtl"> <mrow> <munder> <mi mathcolor="gray"><rl> حد</rl></mi> <mrow> <mi>س</mi> <mo> ← </mo> <mn>2</mn> </mrow> </munder> <mfrac> <mrow> <mi>س</mi> <mo>-</mo> <mn>2</mn> </mrow> <mrow> <mi>س</mi> <mo>-</mo> <mn>10</mn> </mrow> </mfrac> </mrow> </math>	$\underset{\text{س} \leftarrow 2}{\text{حد}} \frac{\text{س} - 2}{\text{س} - 10}$
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <mi>س</mi> <mo rspace="1em">+</mo> <mi>ب</mi> </math>	س + ب

<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <mi>س</mi> <mo>+</mo> <mi>ب</mi> </math>	س + ب
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <ms lquote=';' rquote='>'> الصيغه تكون هنا</ms> </math>	; هنا تكون الصيغه<
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <mrow> <mo lspace='1em' >(</mo> <mn>5</mn> <mo> ، </mo> <mi> س </mi> <mo>) </mo> </mrow> </math>	(5 ، س)
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <mrow> <mo lspace='1em' >(</mo> <mn>5</mn> <mo> , </mo> <mi> س </mi> <mo>) </mo> </mrow> </math>	(5, س)

Table (24): Example of using rspace/lspace and lquote/rquote attributes with RTL.

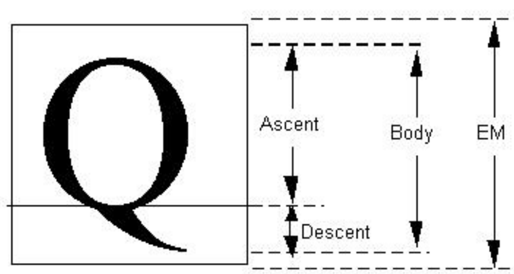


Figure (19): The EM Square (from [29])

4.3.3. Mathematical Fonts

For Arabic mathematical expressions, there are rules for typesetting the notation to get the Arabic mathematical presentation of the document. These rules must respect the features of the specific symbols used and the RTL direction. Further, Arabic notation must follow a homogeneous layout in all documents. Mathematical fonts are used to correctly display MathML for Arabic presentation. There are two fonts used for displaying Arabic expressions:

- ACM (Arabic Computer Modern)
- RamzArab

As previously mentioned, ACM is a font that was introduced in order to have Computer Modern fonts in the TrueType format and to have mirror glyph symbols. A new element, <amath>, was proposed for MathML to use ACM fonts in presenting expressions [12]. But, these fonts still have problems when used for representing Arabic formulas. An <rl> element would need to be included for each subexpression in order to change the direction within each element, requiring heavy markup for Arabic notation. Meanwhile, if <amath> is replaced with <math dir='rtl>, then the ACM font is not used (see Table 25).

MathML Code	Output
<amath xmlns="http://www.w3.org/1998/Math/MathML" > <mtext>نص رياضي</mtext> </amath>	نص رياضي
<math xmlns="http://www.w3.org/1998/Math/MathML" dir='rtl'> <amath> <mtext>نص رياضي</mtext> </amath> </math>	نص رياضي

Table (25): Example of using <amath> for ACM font.

RamzArab was introduced by Cadi Ayyad University in Morocco for Arabic mathematics and is a dynamic font. RamzArab provides a homogeneous Arabic font family in the Open Type format, and respects Arabic calligraphy rules (see Figure 20). RamzArab gives symbols the same shapes, sizes, and boldness [30].

As discussed, there are mathematical symbols in Arabic mathematical handbooks (see the Appendix) that are not yet part of the Unicode standard but for which mirror symbols are available. The RamzArab font includes most common Arabic symbols, which the researchers at Cadi Ayyad University proposed for addition to the Unicode standard [27]. Some of these symbols were added to Unicode version 6.1, under the Private Use Area E000-F8FF of the Basic Multilingual Plane. The RamzArab font can also be used in the TeX system [30].

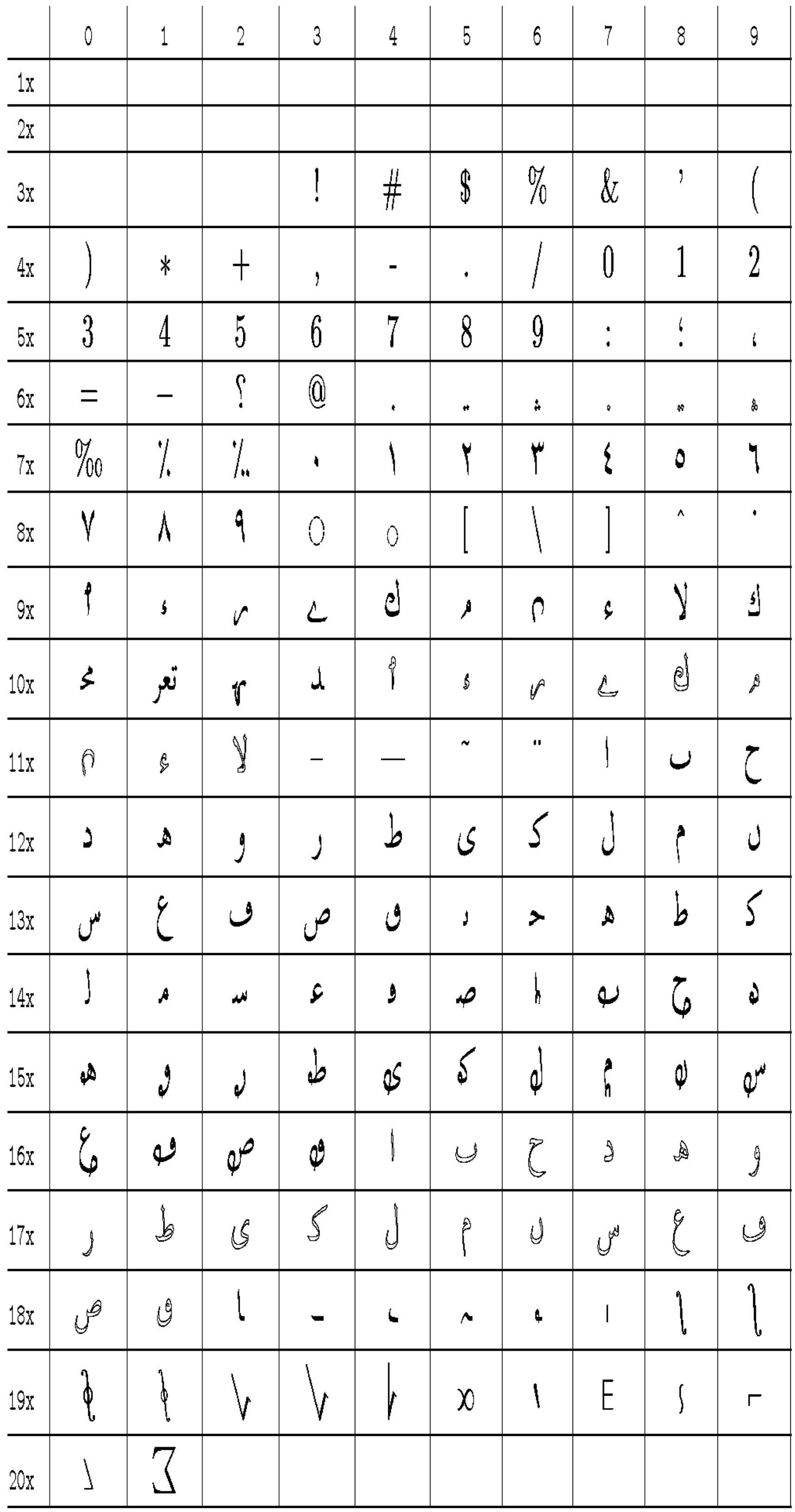


Figure (20): RamzArab Font (from [30])

4.3.4. Kashida (Extensible) Symbols

Arabic writing is smoothed by using joining curves according to calligraphy rules. However, this method is difficult to reproduce when rendering Arabic expressions in software. For instance, the "lim" symbol (for limit) is written with three letters joined and stretched into one word: NON + HAH+ALEF, or . The result is a single ligature (stretched letters of the Arabic alphabet). Unfortunately, this ligature is rendered with a straight line rather than a cursive line (which is required in Arabic calligraphy) [23] (see Figure 21).

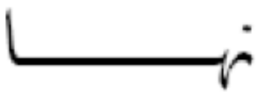
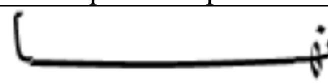
Current Representation	Required Represent
	

Figure (21): Limit Symbol in Arabic Language and the difference between representations. (from [23])

The TeX system allows the curvilinear stretching of Arabic letters (which obeys the calligraphy rules) via the CurExt function and PostScript RamzArab font. CurExt can determine the size of each ligature according to its context, and do horizontal (e.g. limit symbol) and vertical (e.g. parentheses) curvilinear extensions (see Table 26) [30].

TeX with CurExt	MathML
$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

Table (26) : Examples of parentheses in TeX and MathML (first column taken from [12])

5. Conclusion

In this report, we explain the characteristics of Arabic writing and diagnose the issues that appear when attempting to use Arabic character sets in MathML to follow the rules of Arabic calligraphy.

In preparing to create the figures for this report, we found that browser support for Arabic mathematical notation is limited. Only two browsers, Firefox and Dadzilla, support Arabic mathematical notation (Dadzilla is a MathML browser based on the open source Mozilla project). Neither browser supports bi-directionality or cursivity.

It will be necessary to create a bidirectionality algorithm for mixed direction text in MathML (since the Unicode Bidi Algorithm does not consider individual elements separately). Furthermore, there still remain symbols that are not defined in Unicode, including stretched ligatures. Currently, these ligatures are represented by identifiers in <mi> elements without any stretch property.

6. References

[1] PennState "Computing with Accents,Symboles and Forien Scripts " http://symbolcodes.tlt.psu.edu/index.html , Penn State University, 2013.

[2] W3C, "Arabic mathematical notation," http://www.w3.org/TR/2006/NOTE-arabic-math-20060131,W3C Interest Group Note 31 January 2006.

[3] MathML Central, History of MathML "http://www.mathmlcentral.com/history.html", 27 April 2009.

[4] W3C Math Home," http://www.w3.org/Math ", W3C, **15**May, 2009**.**

[5] Yunus Adanur, Oktay,Amed Isik,Mathematics and Language, "http://www.mathnet.or.kr/mathnet/kms_tex/981583.pdf ",Research in Mathematical Education ,March 2004.

[6] Rana Abou Rjeily "Cultural Connectives:Bridging The Latin and Arabi Book", Mark Batty Publisher,New York, 2011

[7] Azzeddine Lazrek,Mohammed Hssini,Mohamed Jamal Eddine Benatia " Arabic Digital Typography design and Connect Characters Issues ", Department of computer Science,University Cadi Ayyad –Marrakech Morocco, 2010.

[8] Richard Ishida , Creating HTML Pages in Arabic, Hebrew and Other Right-to-left Scripts ,"http://www.w3.org/International/tutorials/bidi-xhtml/" ,W3C, 2013.

[9] Richard Ishida , Aharon Lanin ,Inline markup and bidirectional text in HTML , "http://www.w3.org/International/articles/inline-bidi-markup/#usecase1",W3C, 2014

[10] Edo Smitshuijzen "Arabic Font Specimen Book", De Buitenkant,Amsterdam , 2009

[11] Mokhtar Alem Fyyad, A comparative study of the technical attributes in the Thuluth Calligraphy. Supplementary requirement to obtain a Master's degree in Art Education, Faculty of Education in Makkah, Umm Al-Qura University, Saudi Arabia,2001

[12] Mohamed ELYAAKOUBI "Étude et réalisation d'un système de justification du e-document scientifique multilingue", Department of Computer Science, Faculty of Sciences,University Cadi Ayyad - Marrakech, Morocco, Soutenue le 23 octobre 2010

[13] Mohamed Hssini, Azzeddine Lazrek, Mohamed Jamal Eddine Benatia," Diacritics Mark in the Electronic Arabic Document", Department of Computer Science, Faculty of Sciences, University Cadi Ayyad - Marrakech, Morocco.

[14] Mustafa Alzahebi, Azzeddine Lazrek,Khalid Sami " Text encoding tools Arab mathematician and typesetting", Department of Computer Science, Faculty of Sciences,University Cadi Ayyad - Marrakech, Morocco,2008

مصطفى الذاهيبي وعزالدين لزرق وخالد سامي "أدوات ترميز النص الرياضياتي العربي وتنضيده",جامعة القاضي عياض، كلية العلوم، شعبة المعلوماتية,مراكش, المغرب,٢٠٠٨

[15] Rabab Albar, "https://uqu.edu.sa/page/ar/29620 ",The University of Umalquraa, Department of the Mathematics ,Saudi Arabia, Makkah

[16] Azzeddine Lazrek ”Arabic Mathematical Alphabetic Symbols”, Faculty of Sciences, Cadi Ayyad University, Marrakech, Morocco, 2005
[17] J J O'Connor, E F Robertson ,”The Arabic Numeral System “, http://www-history.mcs.st andrews.ac.uk/HistTopics/Arabic_numerals.html, School of Mathematics and Statistics , University of St Andrews, Scotland,January 2001
[18] Drexel University, Ask Dr.Math The Math Forum, http:// mathforum.org /library/ drmath/ view/6132.html , 2002
[19] Planetary, planetMath, “Superscript ” , planetmath.org/superscript , 2013.
[20] John Lawler, “punctuation ”, The International Encyclopedia of Language and Linguistics, University of Michigan ,2nd edition, Elsevier, 2006.
[21] Khalid Sami, “The Needs of the Arab scientific Electronic Documents: The term of the symbolic standard word processing tools to the scientific”, http://dc105.4shared.com/doc/wboftjZt/preview.html , Faculty of Sciences, Cadi Ayyad University, Marrakech, Morocco,2003
خالد سامي. حاجيات الوثائق العلمية العربية الإلكترونية: من المصطلح الرمزي المعياري إلى أدوات معالجة النصوص العلمية. ندوة التعريب في التعليم العالي، دمشق، سوريا، 2003.
[22] Azzeddine Lazrek ”Word Processing Fonts of scientific Arabic”, Department of Computer Science, Faculty of Sciences,University Cadi Ayyad - Marrakech, Morocco.

[23] Abdullah Aldusary,Abdulaziz Alsbaiheen, Fahad Alsweleh,”MathML and Arabic Language”, http://ccisdb.ksu.edu.sa/files/rep6000000.docx, King Saud University, Department of Computer Science, Faculty of Sciences, Riyadh ,Saudi Arabia, 2009
[24] W3C Math Home, “Internationalized, Unicode Bidirectional Algorithm basics”, http://www.w3.org/International/articles/inline-bidi-markup/uba-basics, W3C, 24 April, 2014.
[25] W3C Recommendation,” Presentation Markup”, http://www.w3.org/TR/2014/REC-MathML3-20140410/chapter3-d.html#presm.bidi.token, W3C, 10 April 2014.
[26] Daniel Marquès Solé,” Edition with Arabic mathematical notation”, http://cermat.org/events/MathUI/13/proceedings/Editing-with-Arabic-Notation-MathUI2013.pdf, WIRIS, 14 May 2013.
[27] Azzeddine Lazrek , Khalid Sami,Mohammed Eddahabi ”Arabic Mathematical e-Documents”, Faculty of Sciences, Cadi Ayyad University, Marrakech, Morocco, 2004
[28] Azzeddine LAZREK , “Proposals Symbols for Unicode Consortium”, http://www.ucam.ac.ma/fssm/rydarab/unicode.htm, Faculty of Sciences, Cadi Ayyad University, Marrakech, Morocco,16 March 2008
[29] Osnabrück University, “font and Text “, http://www-lehre.inf.uos.de/~fbstark/diplom/docs/swf/DefineFont.htm , Department of Mathematics, Institute of Computer Science, Marrakech,German.
[30] Azzeddine Lazrek , Khalid Sami, Mohamed ELYAAKOUBI, ”Dynamic Arabic mathematical fonts”, Faculty of Sciences, Cadi Ayyad University, Marrakech, Morocco,2004

7. Appendices

ثانيا – ارسم زاوية موجهة في وضع قياسي بحيث تكون:

أ (٤ ، ٤) ∋ الضلع النهائي.　　ب (-٥ ، -٤) ∋ الضلع النهائي.

جـ (٠ ، ٤) ∋ الضلع النهائي.　　د (-٣ ، ٠) ∋ الضلع النهائي.

Kuwait, Grade 10 (high school), Art Student, Second Term Handbook.

البرهان :

نفرض أن ناتج قسمة د(س) على (س – أ) يساوي ق(س)

∴ د(س) = (س – أ) ق(س) + ر(س) .... (١)

حيث ر(س) هو باقي القسمة .

وحيث إن المقسوم عليه (س – أ) من الدرجة الأولى .

∴ باقي القسمة ر(س) يكون من درجة صفر .

أي أن ر(س) حدودية ثابتة

∵ ر(س) = جـ ، جـ ∈ ح

بوضع س = أ في كل من طرفي المعادلة (١)

∴ د(أ) = صفر × ق(أ) + جـ

∴ جـ = د(أ)

أي أن باقي قسمة د(س) على س – أ هو د(أ)

Kuwait, Grade 10 (high school), Art Student, First Term Handbook.

★ الصورة العامة للحدودية من الدرجة ن في المتغير س هي :

د(س) = أن سن + أن-١ سن-١ + . . . + أ١ س + أ٠ حيث :

أن ، أن-١ ، . . . ، أ١ ، أ٠ ∈ ح ، أن ≠ ٠ ، ن عدد صحيح غير سالب ،

أن يسمى المعامل الرئيسي ، أ٠ يسمى الحد الثابت (الحد المطلق) .

Kuwait, Grade 10 (high school), Art Student, First Term Handbook.

ما سبق :

إذا كان م١ ، م٢ هي ميلا المستقيمين ل١ ، ل٢ على الترتيب فإن :

١ ل١ // ل٢ إذا وفقط إذا م١ = م٢

٢ ل١ ⊥ ل٢ إذا وفقط إذا م١ × م٢ = −١

Kuwait, Grade 10 (high school), Art Student, First Term Handbook.

– تعين ترتيب الوسيط :

ترتيب الوسيط = $\frac{ن}{٢}$

= $\frac{مجـ ت}{٢}$ = ٢٧

Kuwait, Grade 10 (high school), Art Student, First Term Handbook.

وتكون $\overline{بحـ}$ // $\overline{أحـ'}$ .... (I)

من النتيجة (٢) :

∴ ق$\hat{(١)}$ = ق$\hat{(٢)}$ .... (II)

كذلك ، د (ك ، ١٨٠°) :

$\overline{بحـ}$ ⟵ $\overline{ب'أ'}$ .... (III)

وتكون $\overline{بحـ}$ // $\overline{ب'أ'}$

∴ ق$\hat{(٣)}$ = ق$\hat{(٤)}$ .... (IV)

من (I) ، (III) : حـ' ، أ ، ب' على استقامة واحدة

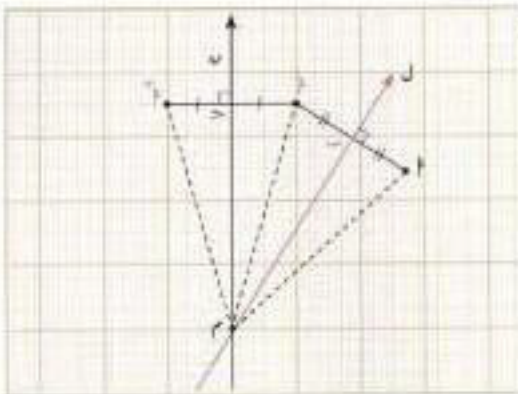
ومن (II) ، (IV) :

ق$\hat{(١)}$ + ق$\hat{(٣)}$ + ق$\hat{(٥)}$ = ق$\hat{(٢)}$ + ق$\hat{(٤)}$ + ق$\hat{(٥)}$ = ١٨٠° ■

Kuwait, Grade 10 (high school), Art Student, First Term Handbook.

انعكاسان متتاليان في مستقيمين متقاطعين يكافئ دورانا مركزه نقطة تقاطع هذين المستقيمين وقياس زاويته يساوي ضعف قياس الزاوية المحددة بهما .

انظر شكل (٤ - ٢٣)



شكل (٤ - ٢٣)

بفرض أن ل ، ه متقاطعان في م ، وبفرض أن أ' صورة أ بالانعكاس في ل ،

أ'' صورة أ' بالانعكاس في ه ، فإن :

م أ = م أ' ، م أ' = م أ'' لماذا؟

∴ م أ'' = م أ' = م أ

كذلك :

ق (أ مٌ أ'') = ٢ ق (ب مٌ حـ) - لماذا؟

∴ أ'' صورة أ تحت تأثير دوران مركزه م في الاتجاه من ل إلى ه وقياس زاويته = ٢ ق (ب مٌ حـ) .

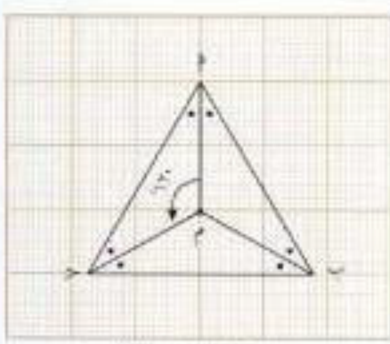
Kuwait, Grade 10 (high school), Art Student, First Term Handbook.

لذلك نقول إن الدائرة متناظرة حول مركزها بدوران قياس زاويته هـ .

كمثال آخر على التناظر الدوراني :

أ ب حـ مثلث متطابق الأضلاع

م نقطة تقاطع منصفات زواياه ، شكل (٤ - ٢٥)



شكل (٤ - ٢٥)

واضح أن : أ م = حـ م = ب م ،

ق (أ مٌ حـ) = ق (حـ مٌ ب) = ق (ب مٌ أ) = ١٢٠°

∴ تحت تأثير د (م ، ١٢٠°)

أ ⟵ حـ ، ب ⟵ أ ، حـ ⟵ ب ، Δ أ ب حـ ⟵ Δ حـ أ ب

أي أن المثلث المتطابق الأضلاع يكون متناظراً تحت تأثير الدوران د (م ، ١٢٠°) حيث م هي

Kuwait, Grade 10 (high school), Art Student, First Term Handbook

في التمارين (10–15)، أوجد معكوس كل دالة مما يلي:

(10) $y=\frac{1}{3}x^3$	**(11)** $y=2\sqrt[4]{x}$	**(12)** $y=\frac{1}{3}x^4$
(13) $y=\frac{1}{3}\sqrt[3]{x}$	**(14)** $y=\sqrt[3]{x-1}$	**(15)** $y=(x+2)^4-3$

Kuwait, Grade 11, Science Student, First Term Handbook.

• يُعطى طول المتجه $\mathbf{v} = \langle v_1, v_2 \rangle$ بالصيغة

$$|\mathbf{v}| = \sqrt{(v_1)^2 + (v_2)^2}$$

Saudi Arabic, Math6, High School, Science Student Handbook p46

53) تحدٍّ: قدّر كلًّا من النهايات الآتية للدالة f إذا كانت موجودة:

$$f(x) = \begin{cases} 2x + 4 , & x < -1 \\ -1 , & -1 \le x \le 0 \\ x^2 , & 1 < x \le 2 \\ x - 3 , & x > 2 \end{cases}$$

a) $\lim_{x \to -1} f(x)$　b) $\lim_{x \to 0} f(x)$　c) $\lim_{x \to 2^+} f(x)$

Saudi Arabic, Math6, High School, Science Student Handbook p138

مفهوم أساسي — **التكامل المحدد**

يُعبر عن مساحة المنطقة المحصورة بين منحنى دالة والمحور x في الفترة $[a, b]$ بالصيغة

$$\int_a^b f(x)dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i)\Delta x , \Delta x = \frac{b-a}{n}, x_i = a + i\Delta x$$

حيث a الحد الأدنى، و b الحد الأعلى، وتُسمَى هذه الطريقة مجموع ريمان الأيمن.

Saudi Arabic, Math6, High School, Science Student Handbook p167

مفهوم أساسي — **النظرية الأساسية في التفاضل والتكامل**

إذا كانت $F(x)$ دالةً أصلية للدالة المتصلة $f(x)$، فإن

$$\int_a^b f(x)\,dx = F(b) - F(a) .$$

ويمكن التعبير عن الطرف الأيمن من هذه العبارة بالرمز $F(x)\Big|_a^b$.

Saudi Arabic, Math6, High School, Science Student Handbook p175

الزاوية بين متجهين	$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{\lvert\mathbf{a}\rvert\lvert\mathbf{b}\rvert}$	مسقط u على v	$\mathbf{w}_1 = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\lvert\mathbf{v}\rvert^2}\right)\mathbf{v}$
طول متجه	$\lvert\mathbf{v}\rvert = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$	الضرب القياسي للثلاثيات	$\mathbf{t}\cdot(\mathbf{u}\times\mathbf{v}) = \begin{vmatrix} t_1 & t_2 & t_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$
الضرب الاتجاهي لمتجهين في الفضاء	$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$		

المشتقات

قاعدة مشتقة القوة	إذا كان $f(x) = x^n$، حيث n عدد حقيقي، فإن $f'(x) = nx^{n-1}$.	قاعدة مشتقة المجموع أو الفرق	إذا كان $f(x) = g(x) \pm h(x)$ فإن $f'(x) = g'(x) \pm h'(x)$.
قاعدة مشتقة الضرب	$\frac{d}{dx}[f(x)\,g(x)] = f'(x)\,g(x) + f(x)\,g'(x)$	قاعدة مشتقة القسمة	$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)\,g(x) - f(x)\,g'(x)}{[g(x)]^2}$

التكاملات

التكامل غير المحدد	$\int f(x)\,dx = F(x) + C$	النظرية الأساسية في التفاضل والتكامل	$\int_a^b f(x)\,dx = F(b) - F(a)$

الإحداثيات القطبية

صيغة القسمة	$\frac{z_1}{z_2} = \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)]$	صيغة الضرب	$z_1z_2 = r_1r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)]$

Saudi Arabic, Math6, High School, Science Student Handbook p186

ص = أ س^{٢} + ب س + جـ ؛ أ ≠ ٠

ب) أوجد القيمة العظمى أو القيمة الصغرى للدالة.

القيمة العظمى هي الإحداثي الصادي للرأس.

الإحداثي السيني للرأس = $\frac{-ب}{٢أ} = \frac{٤}{٢(-٢)} = -١$.

Saudi Arabic, Math3, Intermediate School, Handbook p92

س^{٢} − ٢س − ٨ = ٠	المعادلة الأصلية	س^{٢} − ٢س − ٨ = ٠
(−٢)^{٢} − ٢(−٢) − ٨ ≟ ٠	س = −٢ أو س = ٤	(٤)^{٢} − ٢(٤) − ٨ ≟ ٠

Saudi Arabic, Math3, Intermediate School, Handbook p99

س + ١٢ = ± √٢٦٩٤
س = −١٢ ± √٢٦٩٤

س = −١٢ − √٢٦٩٤
≈ −٦٣,٩

Saudi Arabic, Math3, Intermediate School, Handbook p106

٢٧) أي العبارات الآتية تكافئ √(٣٦/٢٧) :

أ) √(٣/٣) جـ) √٦ / ٣

ب) ٢√٣ / ٣ د) ٣√٢ / ٢

Saudi Arabic, Math3, Intermediate School, Handbook p134

الرموز: جـَ٢ = أَ٢ + بَ٢

Saudi Arabic, Math3, Intermediate School, Handbook p135

قراءة الرياضيات

قياس الزاوية

ق ∠ أ يُقرأ: قياس الزاوية أ.

Saudi Arabic, Math3, Intermediate School, Handbook p146

مثال: إذا كان △ أ ب جـ ∽ △ د هـ و، فإن ق ∠ أ = ق ∠ د، ق ∠ جـ = ق ∠ و،
ق ∠ ب = ق ∠ هـ .
أب / دهـ = جـ ب / هـ و = أجـ / د و = ١ / ٢ .

Saudi Arabic, Math 3, Intermediate School, Handbook p145

$ن\,ق\,ر = \dfrac{ن!}{(ن - ر)!\,ر!}$ قانون التوافيق

Saudi Arabic, Math3, Intermediate School, Handbook p179

$ن\,ل\,ر = \dfrac{ن!}{(ن - ر)!}$ قانون التباديل

Saudi Arabic, Math, Intermediate School, Handbook p180

$$\text{حيث}\quad \frac{\widehat{CD}}{OC} = \frac{\widehat{AB}}{OA} = \frac{\widehat{MN}}{OM}$$

SYRIA, Math 10, High School, Handbook p69

$$\widehat{XZY} = \widehat{KZL} \quad \text{لماذا؟}$$
$$\widehat{YXZ} = \widehat{LKZ}$$
$$\Delta XYZ \sim \Delta KLZ$$

$$\widehat{B} = \widehat{K}\ ،\ \widehat{A} = \widehat{M}\ ،\ \widehat{C} = \widehat{L}$$

SYRIA, Math 10, High School, Handbook p45&47

الرموز الرياضية المستخدمة

ط	مجموعة الاعداد الطبيعية	⊥	عمودى على
ص	مجموعة الأعداد الصحيحة	//	يوازى
ن	مجموعة الأعداد النسبية	$\overline{اب}$	القطعة المستقيمة $\overline{اب}$
نَ	مجموعة الأعداد غير النسبية	$\overleftarrow{اب}$	الشعاع $\overleftarrow{اب}$
ع	مجموعة الأعداد الحقيقية	$\overleftrightarrow{اب}$	المستقيم $\overleftrightarrow{اب}$
$\sqrt{ا}$	الجذر التربيعى للعدد ا	ق (∠ ا)	قياس زاوية ا
$\sqrt[3]{ا}$	الجذر التكعيبى للعدد ا	ق ($\frown{اب}$)	قياس القوس اب
[أ ، ب]	فترة مغلقة	~	تشابه
]أ ، ب[	فترة مفتوحة	<	أكبر من
]أ ، ب]	فترة نصف مفتوحة	⩽	أكبر من أو تساوى
[أ ، ب[	فترة نصف مفتوحة	>	أقل من
[أ ، ∞[	فترة غير محدودة	⩾	أقل من أو تساوى
≡	تطابق	ل(ا)	احتمال وقوع الحدث ا
ن (ا)	عدد عناصر الحدث ا	$\overline{س}$	الوسط الحسابى
ف	فضاء العينة	σ	الانحراف العيارى
		مج أو Σ	المجموع

Egypt, Intermediate School 3 term1, Handbook p7

	1EE0	1EE1	1EE2	1EE3	1EE4	1EE5	1EE6	1EE7	1EE8	1EE9	1EEA	1EEB	1EEC	1EED	1EEE	1EEF
0	1EE00	1EE10		1EE30				1EE70	1EE80	1EE90		1EEB0				1EEF0
1	1EE01	1EE11	1EE21	1EE31		1EE51	1EE61	1EE71	1EE81	1EE91	1EEA1	1EEB1				1EEF1
2	1EE02	1EE12	1EE22	1EE32	1EE42	1EE52	1EE62	1EE72	1EE82	1EE92	1EEA2	1EEB2				
3	1EE03	1EE13							1EE83	1EE93	1EEA3	1EEB3				
4		1EE14	1EE24	1EE34		1EE54	1EE64	1EE74	1EE84	1EE94		1EEB4				
5	1EE05	1EE15		1EE35				1EE75	1EE85	1EE95	1EEA5	1EEB5				
6	1EE06	1EE16		1EE36				1EE76	1EE86	1EE96	1EEA6	1EEB6				
7	1EE07	1EE17	1EE27	1EE37	1EE47	1EE57	1EE67	1EE77	1EE87	1EE97	1EEA7	1EEB7				
8	1EE08	1EE18					1EE68		1EE88	1EE98	1EEA8	1EEB8				
9	1EE09	1EE19	1EE29	1EE39	1EE49	1EE59	1EE69	1EE79	1EE89	1EE99	1EEA9	1EEB9				
A	1EE0A	1EE1A	1EE2A				1EE6A	1EE7A		1EE9A		1EEBA				
B	1EE0B	1EE1B	1EE2B	1EE3B	1EE4B	1EE5B		1EE7B	1EE8B	1EE9B	1EEAB	1EEBB				
C	1EE0C	1EE1C	1EE2C				1EE6C	1EE7C	1EE8C		1EEAC					
D	1EE0D	1EE1D	1EE2D		1EE4D	1EE5D	1EE6D		1EE8D		1EEAD					
E	1EE0E	1EE1E	1EE2E		1EE4E		1EE6E	1EE7E	1EE8E		1EEAE					
F	1EE0F	1EE1F	1EE2F		1EE4F	1EE5F	1EE6F		1EE8F		1EEAF					

The Arabic Mathematical Alphabetic Symbols in Unicode 6.1
http://www.unicode.org/charts/PDF/Unicode-6.1/U61-1EE00.pdf

From: Azzeddine LAZREK Azzeddine LAZREK <lazrek@uca.ma>
Subject: Re: Need your kind help
Date: 10 March, 2014 2:32:09 AM EDT
To: Mona Alshehri <malshehr@uwo.ca>

1 Attachment, 9 KB

ليس كذلك، فمن أهم جديد الاصدار الثالثة من الماث ام ال هو دعم الصيغة الرياضياتية العربية الخالصة أي ما سميتها بالمشرقية باستعمال:

dir="rtl"

وهذا النمط ليس فقط "يختلف عن النمط المغربي بانه لايستخدم الارقام الاتينيه" كما قلتي.

2014-03-10 0:17 GMT+00:00 Mona Alshehri <malshehr@uwo.ca>:

> اشكرك جزيل الشكر استاذي الفاضل
>
> بالفعل لقد قرآت بحثك عن هذه المشاكل
> ولقد صادفت سوال اخر اثناء تصفح الماث ام ال وهو ان كتابة المعادلات الرياضيه بالنمط المشرقي لم يتم تعريفه
> وهذا النمط هو كتابة المعادلات الرياضيه والاحرف والرموز والارقام بالنص العربي ويختلف عن النمط المغربي بانه لايستخدم الارقام الاتينيه
>
> لماذا لم يتم تضمين هذا النمط وماهي المشاكل التي واجهوها لعدم تمكنهم من تضمينه؟
>
> $$\text{ت}(\text{س}) = \begin{cases} \underset{\text{ب}=١}{\overset{\text{ص}}{\text{مجـ}}}\ \text{س}^{\text{ب}} & \text{إذا كان } \text{س} > ٠ \\ \int_{١}^{\text{ص}} \text{س}^{\text{ب}}\ \text{ء}\text{س} & \text{إذا كان } \text{س} \in \text{م} \\ \text{ظا}\ \pi & \text{غير ذلك } (\text{مع } \pi \simeq ٣٫١٤١) \end{cases}$$
>
> Thanks & Regards,
>
> Mona Alshehri
> malshehr@uwo.ca
> University of Western Ontario
> London ON, Canada
> Master Of Software Engineering
> Mobile # +1(519)701-5977
>
> On 2014-03-09, at 8:08 PM, Azzeddine LAZREK Azzeddine LAZREK wrote:
>
>> بالنسبة للتيبوغرافية، من أهم الإشكاليات هو الاتجاه، خصوصا عند تواجد نصين باتجاهين مختلفين ؛ ثم محاذاة النص مع مراعاة قواعد الخط العربي ؛ ثم وضع الحركات دون لبس.
>>
>> مع المناقشة تظهر الأشياء...
>>
>> بالتوفيق والسداد،
>>
>> عزالدين
>>
>> 2014-03-09 0:37 GMT+00:00 Mona Alshehri <malshehr@uwo.ca>:
>>
>>> انا ممتنه جدا لمساعدتك لي
>>> لقد طلب مني ان اقوم بتحديد المشاكل التي نواجهها كمستخدمين للغة العربيه في كتابة المعادلات الرياضيه اثناء استخدام البرامج
>>> على اجهزة الحاسب وماهي ابرز هذه المشاكل التي تواجهنا ، وتحليلها من ناحية حاسوبية وجبريه
>>> لقد وجدت مجموعة من هذه المشاكل التي قمتم بذكرها في بعض دراساتكم وذكرتها بالاسفل
>>> سوالي هو هل بامكانكم مساعدتي بذكر بعضامن ابرز المشاكل النتي نواجهها عندما نريد استخدام اللغه العربيه على آجهزة الحاسب ؟
>>>
>>> Thanks & Regards,

Thanks & Regards,

Mona Alshehri
malshehr@uwo.ca
University of Western Ontario
London ON, Canada
Master Of Software Engineering
Mobile # +1(519)701-5977

On 2014-03-08, at 7:28 PM, Azzeddine LAZREK Azzeddine LAZREK wrote:

Wa ealaykum salam,

I have many studies about Arabic mathematical notation, some years ago.

However, I don't understand your need.

If you can explain more, perhaps in Arabic this first time, I'll help you with pleasure.

Best regards,
Azzeddine

2014-03-09 0:02 GMT+00:00 Mona Alshehri <malshehr@uwo.ca>:

Dear Dr.Azzeddine

Alsalamu Alikum,,,,,

My name is Mona from Saudi Arabia and student at the University of Western Ontario with Doctor Stephen Watt. I am doing my thesis on the issues of rendering Mathematical Expression or Notation in the Computer Program.

I did some search to collect and analysis some issues but can you help me and give me the major issue or a list of the issues that we face as Arabic end users for programmes who are using Mathematical notations. the issues I found are:

- The limitation of mathematical font families that are used to provide homogeneous font.
- Extension, Kashida and stretching symbols in Arabic expressionà curve line not straight one.(such as horizontal and vertical stretching)
- There are many symbols in Arabic mathematical expression, found in Arabic books, which are not encoding in the Unicode Standard.

§ example (units ànot included in the Unicode Standard)

- Variable in Arabic notation is written by using isolated letters and ligatures as well.
- Superscript and subscript appear on the left instead of right
- Mirroring for Arabic mathematical symbols are not available.
- The meaning of some mathematical symbols is according to the context

did I miss something or can I find more issues that I can study in this area?

Thanks & Regards,

Mona Alshehri
malshehr@uwo.ca
University of Western Ontario
London ON, Canada
Master Of Computer Algebra