Others Symbols,
Additional characters proposed to Unicode

Azzeddine LAZREK

`lazrek@ucam.ac.ma`

Cadi Ayyad University, Faculty of Sciences

P.O. Box 2390, Marrakech, Morocco

Phone: +212 44 43 46 49 Fax: +212 44 43 74 09

URL: `http://www.ucam.ac.ma/fssm/rydarab`

March 26, 2006

Abstract

Here are gathered the symbols which are not accepted for inclusion to the Unicode Standard during the discussions.

Contents

1 Mirrored symbols

The majority of symbols used in Arabic mathematical presentation are the mirrored corresponding used in Latin presentation. Some symbols and there mirrored one are used together in Latin presentation [5] [3].

In the Unicode Standard, there is the mirrored propriety for some characters [4]. Then, characters that have the mirrored property can be displayed, through an available font and rendering engine, with a mirrored image in right-to-left text runs [1].

1.1 Frequently used symbols

In Arabic Mathematical notation [2], they are some frequently used Arabic symbols that have appropriate mirrored characters in the Unicode Standard (see Table 1). As, the number of those characters is not very important, many characters and their mirrored (e.g., < and >) are already both encoded and the presence of some names ambiguity, we propose them for addition to the Unicode Standard. That allows to obtained them directly.

    ⅃   REVERSED COMPLEMENT
         ≈ <reversed> 2201 ∁

    ƍ   REVERSED PARTIAL DIFFERENTIAL
         ≈ <reversed> 2202 ∂

    E   REVERSED THERE EXISTS
         ≈ <reversed> 2203 ∃

    ⅀   REVERSED N-ARY SUMMARY
         ≈ <reversed> 2211 Σ

    √   REVERSED SQUARE ROOT
         ≈ <reversed> 221A √

    ⌐   REVERSED RIGHT ANGLE
         ≈ <reversed> 221F ∟

    ⦣   REVERSED ANGLE
         ≈ <reversed> 2220 ∠

    ⦡   REVERSED MEASURED ANGLE
         ≈ <reversed> 2221 ∡

    ⦣   REVERSED SPHERICAL ANGLE
         ≈ <reversed> 2222 ∢

    ∫   REVERSED INTEGRAL
         ≈ <reversed> 222B ∫

Table 1: Frequently used symbols with appropriate mirrored image

1.2 Non frequently used symbols

The other symbols are not frequently in use. Of course, the corresponding regular character has the mirrored property, and then characters will be displayed with a mirrored image in right-to-left text runs (see example in Table 2).

⨏⨏	REVERSED DOUBLE INTEGRAL	
	≈ <reversed> 222C ∬	
⨑⨑⨑	REVERSED TRIPLE INTEGRAL	
	≈ <reversed> 222D ∭	
∮	REVERSED CONTOUR INTEGRAL	
	≈ <reversed> 222E ∮	

Table 2: Non frequently used symbols with appropriate mirrored image

1.3 Negated symbols

The slanted line for negation of operators is oriented in the Arabic mathematical presentation as it is for the Latin case. Care has to be taken in mirroring the negated symbols (see example in Table 3). In a similar way, the arrows in symbols like 2231 to 2233 are not mirrored as their orientation is not based on the writing direction, but on the direction in which the integration is applied.

∄	REVERSED THERE DOES NOT EXISTS	
	≈ <reversed> 2204 ∄	

Table 3: Negation symbol

2   Radix symbols with Arabic digits

The radix symbols with Arabic digits from Table 4 are used in Arabic mathematical context. While mathematical layout software can use markup to create radix symbols of any order, these two are common enough in general usage that explicit code points should be assigned. Consequently they are proposed for encoding here.

These characters should have general category Sm, neutral right-to-left directionality and should not mirror.

$\sqrt[3]{}$	REVERSED CUBE ROOT
	≈ \<reversed> 221B $\sqrt[3]{}$
$\sqrt[4]{}$	REVERSED FOURTH ROOT
	≈ \<reversed> 221C $\sqrt[4]{}$

Table 4: Mathematical symbols with no appropriate mirroring

3   Crescent

The crescent sign in the two direction (see Table 5) is proposed for encoding here.

This character should have general category So and have neutral directionality.

⟅	LEFT CRESCENT
⟆	RIGHT CRESCENT

Table 5: Crescent sign

4 Combined symbols

In Latin mathematical notation, the "equal by definition" $\overset{\mathrm{def}}{=}$ operator symbol is already existing (see Table 6). The character "equivalent by definition" $\overset{\mathrm{def}}{\Longleftrightarrow}$ is frequently used in mathematics. The "combining definition" $\overset{\mathrm{def}}{\bigcirc}$ is needed in order to be able to use it with another character (see Table 7).

The two characters COMBINING DEFINITION $\overset{\mathrm{def}}{\bigcirc}$ and EQUIVALENT TO BY DEFINITION $\overset{\mathrm{def}}{\Longleftrightarrow}$ are proposed to be add in Unicode [3].

$\overset{\mathrm{def}}{=}$	EQUAL TO BY DEFINITION

Table 6: Existing by definition symbol

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

Table 7: Proposal by definition symbols

In Arabic mathematical notation, the Arabic operators "equal to by definition" and "equivalent to by definition" are proposed as either compact symbols and element that can be combined with other symbols (see Figure 16 and Table 8).

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

Table 8: Arabic by definition symbols

5 Dates

Two main calendars are in use in the Islamic cultural area: the Hejry (Islamic calendar) and the MylAdy (Gregorian calendar). As the famous marks `AC.` and `BC.` used to distinguish years after and before the year zero. The marks هـ and م are used to make a distinction between the two calendars. Of course, these signs may be added (see Table 9) as special characters.

ه‍ ARABIC HEJRY DATE
  ≈ <font> <isolated> 0647 ه Arabic letter heh
  ≈ <font> 0647 ه Arabic letter heh
  ≈ <font> 06BE ه Arabic letter heh doachashmee
  ≈ <font> FEEB ه Arabic letter heh initial form

م ARABIC MYLADY DATE
  ≈ <font> <isolated> 0645 م Arabic letter meem
  ≈ <font> FEE1 م Arabic letter meem isolated form

Table 9: Dates marks

6 Units

The Arabic square unit, corresponding to the Latin square Km is marked with a special abbreviation كم . It represents an example of various signs for units that are worth of being included in the Unicode Standard (see Table 10).

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

Table 10: Example of unit symbols

7 Arrows

The signs listed in the range U+2790-27FF don't include symmetrical signs oriented right-to-left. So, supplemental arrows in Dingbat arrows should be added even if that the majority of these characters is decorative. In particular, symmetrical signs for all symbols from 2794 to 27BE (see Table 11) are proposed with the specification LEFTWARDS, in contrast to RIGHT-WARDS. Those characters need to be mirrored, but would not be mirrored in plain text since arrows don't have the Bidi Mirrored property in Unicode.

← TRIANGLE-HEADED LEFTWARDS ARROW
← HEAVY TRIANGLE-HEADED LEFTWARDS ARROW

Table 11: Examples of needed arrows

8 Alphabetic symbols

The list of character of *Arabic mathematical alphabetic symbols*, already present in Unicode, proposed to be add into the Unicode Standard is presented below:

- Arabic mathematical isolated alphabetic symbols (see Table 12);

- Arabic mathematical initial alphabetic symbols (see Table 13);

- Arabic mathematical others alphabetic symbols (see Table 14).

From the Unicode and WG2 perspective, any proposal that contains shapes that appear identical to existing characters will not be easy. The problem with duplicating characters with identical shapes is that users will end up entering the wrong character and not be able to determine which is which by looking at the text. The kinds of overriding reasons that Unicode or WG2 would allow is very limited. The initial and isolated forms, as well as to some other characters, such as lam-alef, peh, the hamza and all the other characters from existing characters, used in mathematics are identical in appearance to the characters in the standard Arabic or the Presentation forms blocks. In particular, the presentation forms, are not supposed to be used as character codes for ordinary text. Ordinary text uses the codes for the abstract forms in the 0600 block exclusively, and any positional shapes are determined by the rendering engine. Therefore, not adding another set of mathematical presentation forms will not affect ordinary text users and it's much simpler.

ا MATHEMATICAL ALEF
≈ <font> 0627 ا Arabic letter alef

ب MATHEMATICAL BEH
≈ <font> 0628 ب Arabic letter beh

ج MATHEMATICAL JEEM
≈ <font> 062C ج Arabic letter jeem

د MATHEMATICAL DAL
≈ <font> 062F د Arabic letter dal

ه MATHEMATICAL HEH
≈ <font> 0647 ه Arabic letter heh

و MATHEMATICAL WAW
≈ <font> 0648 و Arabic letter waw

ز MATHEMATICAL ZAIN
≈ <font> 0632 ز Arabic letter zain

ح MATHEMATICAL HAH
≈ <font> 062D ح Arabic letter hah

ط MATHEMATICAL TAH
≈ <font> 0637 ط Arabic letter tah

ي MATHEMATICAL YEH
≈ <font> 064A ي Arabic letter yeh

ك MATHEMATICAL KAF
≈ <font> 0643 ك Arabic letter kaf

ل MATHEMATICAL LAM
≈ <font> 0644 ل Arabic letter lam

م MATHEMATICAL MEEM
≈ <font> 0645 م Arabic letter meem

ن MATHEMATICAL NOON
≈ <font> 0646 ن Arabic letter noon

س MATHEMATICAL SEEN
≈ <font> 0634 س Arabic letter seen

ع MATHEMATICAL AIN
≈ <font> 0639 ع Arabic letter ain

ف MATHEMATICAL FEH
≈ <font> 0641 ف Arabic letter feh

ص MATHEMATICAL SAD
≈ <font> 0635 ص Arabic letter sad

ق MATHEMATICAL QAF
≈ <font> 0642 ق Arabic letter qaf

ر MATHEMATICAL REH
≈ <font> 0631 ر Arabic letter reh

ش MATHEMATICAL SHEEN
≈ <font> 0634 ش Arabic letter sheen

ت MATHEMATICAL TEH
≈ <font> 062A ت Arabic letter teh

ث MATHEMATICAL THEH
≈ <font> 062B ث Arabic letter theh

خ MATHEMATICAL KHAH
≈ <font> 062E خ Arabic letter khah

ذ MATHEMATICAL THAL
≈ <font> 0630 ذ Arabic letter thal

ض MATHEMATICAL DAD
≈ <font> 0636 ض Arabic letter dad

ظ MATHEMATICAL ZAH
≈ <font> 0638 ظ Arabic letter zah

غ MATHEMATICAL GHAIN
≈ <font> 063A غ Arabic letter ghain

Table 12: Mathematical isolated alphabetic symbols

ب MATHEMATICAL BEH
≈ <font> <initial> 0628 ب Arabic letter beh
≈ <font> FE91 ب Arabic letter beh initial form

ج MATHEMATICAL INITIAL JEEM
≈ <font> <initial> 062C ج Arabic letter jeem
≈ <font> FE9F ج Arabic letter jeem initial form

ه MATHEMATICAL INITIAL HEH
≈ <font> <initial> 0647 ه Arabic letter heh
≈ <font> FEEB ه Arabic letter heh initial form

ح MATHEMATICAL INITIAL HAH
≈ <font> <initial> 062D ح Arabic letter hah
≈ <font> FEA3 ح Arabic letter hah initial form

ي MATHEMATICAL INITIAL YEH
≈ <font> <initial> 064A ي Arabic letter yeh
≈ <font> FEF3 ي Arabic letter yeh initial form

ك MATHEMATICAL INITIAL KAF
≈ <font> <initial> 0643 ك Arabic letter kaf
≈ <font> FEDB ك Arabic letter kaf initial form

ل MATHEMATICAL INITIAL LAM
≈ <font> <initial> 0644 ل Arabic letter lam
≈ <font> FEDF ل Arabic letter lam initial form

م MATHEMATICAL INITIAL MEEM
≈ <font> <initial> 0645 م Arabic letter meem
≈ <font> FEE3 م Arabic letter meem initial form

ن MATHEMATICAL INITIAL NOON
≈ <font> <initial> 0646 ن Arabic letter noon
≈ <font> FEE7 ن Arabic letter noon initial form

س MATHEMATICAL INITIAL SEEN
≈ <font> <initial> 0634 س Arabic letter seen
≈ <font> FEB3 س Arabic letter seen initial form

ع MATHEMATICAL INITIAL AIN
≈ <font> <initial> 0639 ع Arabic letter ain
≈ <font> FECB ع Arabic letter ain initial form

ف MATHEMATICAL INITIAL FEH
≈ <font> <initial> 0641 ف Arabic letter feh
≈ <font> FED3 ف Arabic letter feh initial form

ص MATHEMATICAL INITIAL SAD
≈ <font> <initial> 0635 ص Arabic letter sad
≈ <font> FEBB ص Arabic letter sad initial form

ق MATHEMATICAL INITIAL QAF
≈ <font> <initial> 0642 ق Arabic letter qaf
≈ <font> FED7 ق Arabic letter qaf initial form

ش MATHEMATICAL INITIAL SHEEN
≈ <font> <initial> 0634 ش Arabic letter sheen
≈ <font> FEB7 ش Arabic letter sheen initial form

ت MATHEMATICAL INITIAL TEH
≈ <font> <initial> 062A ت Arabic letter teh
≈ <font> FE97 ت Arabic letter teh initial form

ث MATHEMATICAL INITIAL THEH
≈ <font> <initial> 062B ث Arabic letter theh
≈ <font> FE9B ث Arabic letter theh initial form

خ MATHEMATICAL INITIAL KHAH
≈ <font> <initial> 062E خ Arabic letter khah
≈ <font> FEA7 خ Arabic letter khah initial form

ض MATHEMATICAL INITIAL DAD
≈ <font> <initial> 0636 ض Arabic letter dad
≈ <font> FEBF ض Arabic letter dad initial form

غ MATHEMATICAL INITIAL GHAIN
≈ <font> <initial> 063A غ Arabic letter ghain
≈ <font> FECF غ Arabic letter ghain initial form

Table 13: Mathematical initial alphabetic symbols

ﮮ MATHEMATICAL YEH RAJIAT

≈ <font> 06D2 ﮯ Arabic letter yeh barree

ﯕ MATHEMATICAL KAF ZIDANY

≈ <font> 06AA ﯕ Arabic letter swash kaf

ﻻ MATHEMATICAL LAMALEF

≈ <font> FEFB ﻻ Arabic ligature lam with alef

ﻯ MATHEMATICAL ALEF MAKSURA

≈ <font> 0649 ﻯ Arabic letter alef maksura

ﺀ MATHEMATICAL HAMZA

≈ <font> 0621 ﺀ Arabic letter hamza

ﺐ MATHEMATICAL DOTLESS BEH

≈ <font> 066E ﺐ Arabic letter dotless beh

ﺪ MATHEMATICAL INITIAL DOTLESS BEH

≈ <font> FBE8 ﺪ Arabic letter Uighur Kazakh Kirghiz alef maksura initial form

ﻮ MATHEMATICAL DOTLESS FEH

≈ <font> 066F ﻮ Arabic letter dotless feh

ﻮ MATHEMATICAL INITIAL DOTLESS FEH

ﭖ MATHEMATICAL PEH

≈ <font> 067E ﭖ Arabic letter peh

ﭺ MATHEMATICAL TCHEH

≈ <font> 0686 ﭺ Arabic letter tcheh

ﭪ MATHEMATICAL VEH

≈ <font> 06A4 ﭪ Arabic letter veh

ﻉ MATHEMATICAL GHEH

≈ <font> 06A0 ﻉ Arabic letter ain with tree dots above

Table 14: Mathematical others alphabetic symbols

References

[1] Unicode Technical Report ♯25, *Unicode Support for Mathematics*, http://www.unicode.org/reports/tr25/.

[2] W3C Math Interest Group Note, *Arabic Mathematical Notation*, http://www.w3.org/TR/arabic-math/.

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

[4] Unicode, *Bidi Mirroring Glyph Property*, http://www.unicode.org/Public/UNIDATA/BidiMirroring.txt.

[5] Mohamed Jamal Eddine Benatia, Azzeddine Lazrek and Khalid Sami, *Arabic mathematical symbols in Unicode*, Internationalization and Unicode Conference (IUC), IUC 27, Berlin, Germany, April 6-8, 2005, http://www.ucam.ac.ma/fssm/rydarab/doc/communic/unicodem.pdf.
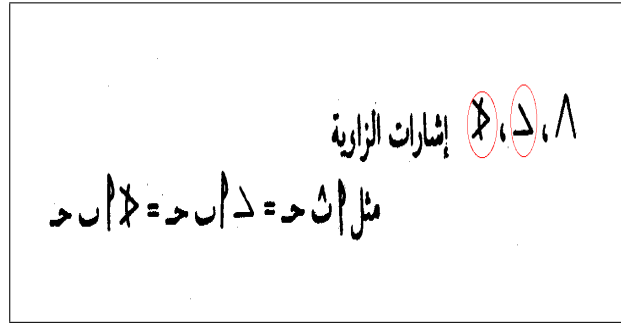
Figure 1: Angle symbol in Amman Convention [1.1]
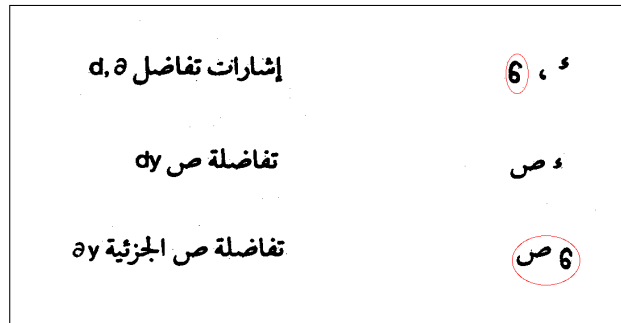
Figure 2: Differential symbol in Amman Convention [1.1]
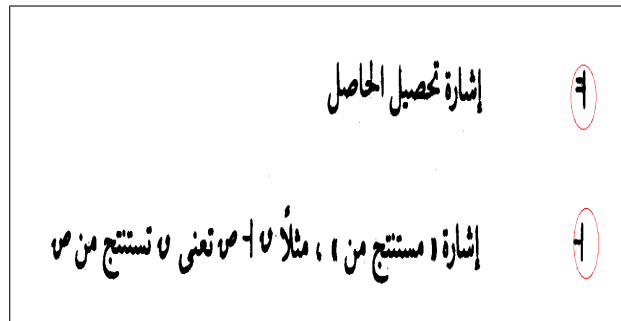
Figure 3: Consequence symbol in Amman Convention [1.1]
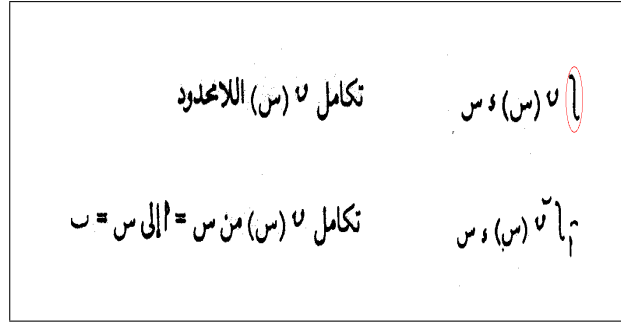
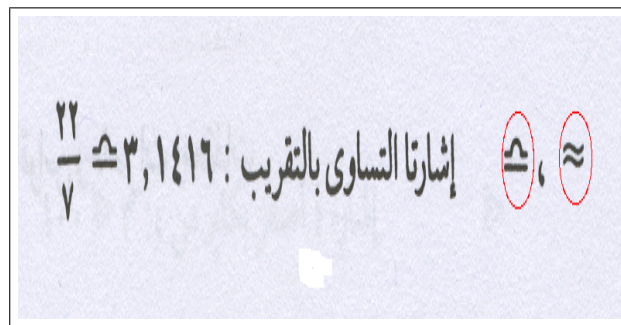Figure 4: Integral symbol in Amman Convention [1.1]

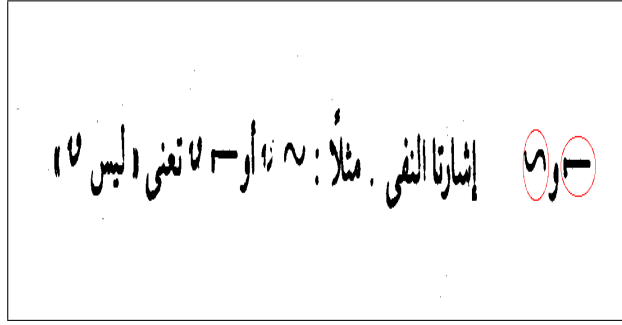Figure 5: Asymptotically equal to symbol in Amman Convention [1.1]

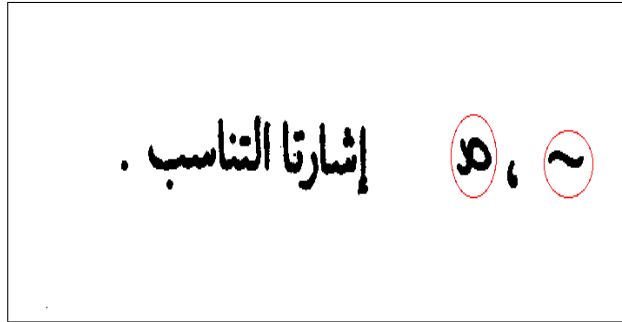Figure 6: Negation symbol in Amman Convention [1.1]

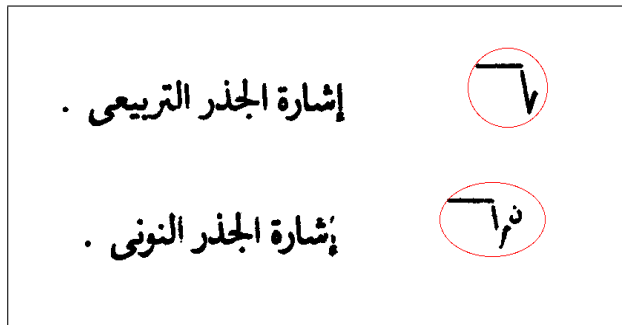Figure 7: Proportional symbol in Amman Convention [1.1]
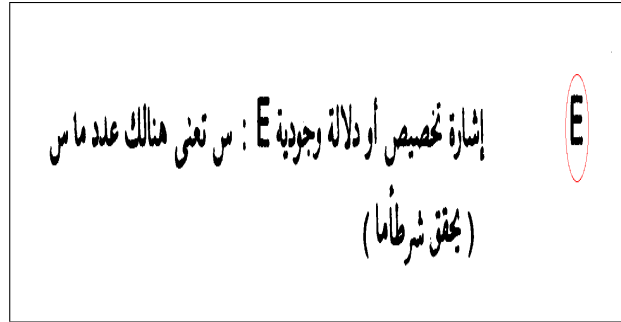

Figure 8: Root symbol in Amman Convention [1.1]

14

Figure 9: There exists symbol in Amman convention [1.1]

Figure 10: Some symbols in Handbook

15

فكلما تغيرت السرعة والزمن تتغير تبعاً لهما المسافة المقطوعة ، ولذا يقال أن :

$$\text{ف} = \text{د} (\text{ع} ، \text{ن})$$

ومساحة شبه المنحرف م الذى أطوال قاعدتيه وارتفاعه هى على الترتيب
قَ ، قّ ، ع نحصل عليها من الصورة :

$$\text{م} = \frac{(\text{قَ} + \text{قّ})}{2} \times \text{ع} .$$

أى أن :

$$\text{م} = \text{د} (\text{قَ} ، \text{فّ} ، \text{ع}) .$$

Figure 11: Prime symbol in Handbook [3.5]

نصف القطر ‏ ‏ ‏ ‏ ‏ ‏ ‏ ‏ ‏ ‏ سه

Figure 12: Ray symbol in Amman convention [1.1]

كذلك إذا كان إم = سه فإنه يمكن كتابة هذه العلاقة بصيغة أخرى هكذا

$$\text{لو}_{\text{إ}} \text{ سه} = \text{م}$$

Figure 13: Logarithm symbol in Handbook [3.3]

16

عدد توفيقـات هـ عنصرا من مجموعة تشمل ه عنصرا هو :

$$_هـق^ه = \frac{ه\,(ه-1)\,\ldots\,(ه-هـ+1)}{هـ\,!} \qquad \text{اذا كان} \; 1 \, \langle \, هـ \, \langle \, ه$$

$$_هـق^ه = 1$$

$$_هـق^هـ = 0 \qquad \text{اذا كان} \; ه \, \langle \, هـ$$

Figure 14: Symbol in Handbook [3.12]

عدد ترتيبات هـ عنصرا من مجموعة عدد عناصرها ه عنصرا هو :

$$_هـت^ه = (ه-1)\,\ldots\,(ه-هـ+1) \qquad \text{اذا كان} \; ه \, \langle \, هـ$$

$$_هـت^ه = 0 \qquad \text{اذا كان} \; ه \, \langle \, هـ$$

Figure 15: symbol in Handbook [3.12]
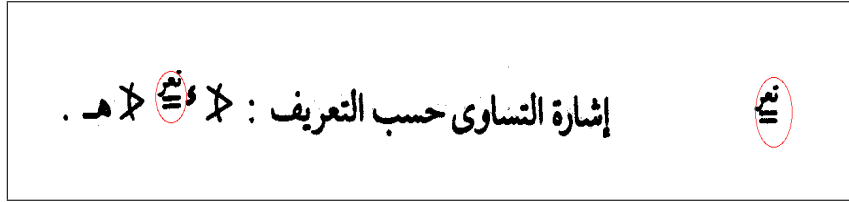
إشارة التساوى حسب التعريف : ⋨ اثہ ⋨ هـ .

Figure 16: Equal by definition symbol in Amman Convention [1.1]