Doc Type: Working Group Document

Title: Raqm Numerals: A Model for Encoding the Siyaq System of South Asia

Source: Anshuman Pandey (pandey@umich.edu)

Status: Individual Contribution **Action:** For consideration by UTC

Date: 2009-04-21

1 Introduction

The intent of this document is to determine models for encoding numerals of the Siyaq notation system in the Universal Character Set (ISO/IEC 10646). It does so through an analysis of the Raqm Numerals, the Siyaq sub-system used in South Asia.

This document draws upon information originally presented in L2/07-414 "Proposal to Encode Siyaq Numerals in ISO/IEC 10646" (December 2007). In L2/07-414, the present author analyzed the four Siyaq traditions and recommended a unified encoding for the numerals of these sub-systems. Further research has indicated that although the numerals of the Diwani, Ottoman, Persian, and South Asian traditions are based upon a common pattern, there are sufficient differences in character typology and orthography to warrant the independent encoding for the numerals of each sub-system.

This document is intended to stand in comparison to L2/09-140 "Diwani Numerals: Towards a Model for Encoding Numerals of the Siyaq Systems" (April 2009). In L2/09-140, the present author described the Diwani Numerals, which possesses the smallest character repertoire of the four sub-systems and the least technical requirements for shaping and other rendering behaviors. The presentation here of the Raqm system may be compared with that of the Diwani Numerals in order to understand not only the differences in the character typology of the numerals of the Siyaq family, but also the locale-specific linguistic factors that differentiates the Siyaq sub-systems.

2 Background

In South Asia, the Siyaq system is known as Raqm (Arabic رقم raqm 'account'). Similar to the other Siyaq traditions, the Raqm Numerals are a specialized subset of the Arabic script that was used for accounting and other numerical notation. The basic Raqm Numerals are stylized monograms of the Arabic names for the numbers, but the numerals for large decimal orders are derived not from Arabic, but from Indic languages, and the method of writing fractions and currencies is based on a common north Indic numerical notation system.

While the majority of documents containing Siyaq are hand-written, a rare exception is the appearance of Raqm Numerals in printed books. A work by Francis Gladwin titled *A Compendious System of Bengal Revenue Accounts* (1790) is perhaps the first book to use Siyaq-based numerals in print. In the preface to his book, Gladwin writes "that the following compendium of Siyak Accounts is the first specimen of the sort that has yet appeared in print, the types having been made purposely for it." A specimen of Raqm printed in Gladwin's metal fonts is given in Figure 9.

¹Gladwin 1790: vii.

	<i>x</i> 1	<i>x</i> 10	<i>x</i> 100	<i>x</i> 1,000	<i>x</i> 10,000	<i>x</i> 100,000	<i>x</i> 1,000,000
1	عم (له)	(حد) هد	٦	الستة	عتث	ديمكه	يلك
2	(سد) لهد	(سع) عد	N	عے	عت	040	عالك
3	(_)	_	_			يدىك	_
4	الله (العد)					ىلول <i>ىك</i>	
5	()					<i>وبک</i>	
6	(-)	ــه				لي مك	
7	معہ (معت	موه	ប	موسية	سوت	سى مىك	سولک
8	(-) L	(ك) ما				مامک	
9	لع (لع)	لحسه	نعا	ل <u>ع_ =</u>	ىرىپ	<i>لولک</i>	<i>يونک</i>

Table 1: South Asian forms of the Siyaq numerals for six decimal orders.

3 The Notation System

3.1 Structure

Raqm Numerals represent units of a base-10 (decimal) positional system. The notation system is additive, that is, the value of a number is the sum of the values of the numerals that constitute it. There is no character for zero; it is inherently represented in the distinct numerals for the various decimal orders.

3.2 Directionality

Raqm Numerals are written right-to-left in the regular manner of the Arabic script, unlike the left-to-right directionality of the Arabic-Indic digits.

3.3 Typology

Raqm Numerals are highly stylized monograms of the Arabic names for numbers for the primary units and their magnitudes in the orders of tens, hundreds, thousands and ten thousands. Raqm Numerals for the hundred thousands and millions are based upon names for numbers derived from Sanskrit.

The numerals may be decomposed to some degree into basic forms and into distinctive signs for the various decimal orders. The exceptions are generally the numerals for magnitudes of one (1, 10, 100, etc.) and two (2, 20, 200, etc.), which have unique forms in all decimal orders. A complete description of the numerals of each decimal order is given in section 4, however, a summary is given below:

- The independent shape of a primary numeral can be decomposed into a character primitive that forms the basis for numerals of all decimal orders for that numeral: FIVE PRIMITIVE

 FIVE.
- Each primary numeral has a base form that is used in writing composite numbers. The base form of a primary numeral is produced by joining the primitive form of the primary numeral to a horizontal stroke: BASE BASE MARK + FIVE PRIMITIVE.
- The numerals for the tens consist of the base form of the primary numeral joined to a distinctive terminal: ← FIFTY ← ← TENS MARKS + ← FIVE BASE.
- The numerals for the hundreds consist of the base form of the primary numeral joined to the numeral for 100:

 FIVE HUNDRED ←

 ONE HUNDRED +

 FIVE BASE. Certain numerals for the hundreds use an alternate base form.
- The numerals for the thousands consist of the primitive form of the primary numeral joined to a distinctive terminal: FIVE THOUSAND THOUSANDS MARK + FIVE PRIMITIVE. When written independently, the thousands are marked with the placeholder , eg. FIVE THOUSAND
- The numerals for the ten thousands consist of a modified (base) form of the tens numerals, which is produced by dropping the TENS MARKS, and a placeholder when written independently: FIFTY THOUSAND PLACEHOLDER + FIFTY BASE. For most tens numeral, the base is identical to the base form of the primary numeral, but different for TEN, TWENTY, and EIGHTY, eg. the base form of EIGHTY is and the base form of EIGHT is —.
- The numerals for the hundred thousands are written using the independent form of the primary numeral and a mark that represents the hundred thousands unit:

 → FIVE HUNDRED THOUSAND ← → FIVE
- The numerals for the millions are written using a modified (base) form of the tens (similar to the ten thousands) and the mark that represents the hundred thousands unit:

 FIVE MILLION

 HUNDRED THOUSANDS MARK +

 FIFTY BASE
- The numerals for the ten millions are written using the independent form of the primary numeral and a mark that represents the ten millions unit:

 TEN MILLION + TEN MILLION + FIVE

3.4 Ordering

The ordering of Raqm Numerals is visual, which reflects the method of expressing numbers in Arabic. The ordering of Raqm Numerals based upon non-Arabic sources is also visual, which reflects the method of expressing numbers in Urdu.

3.5 Orthography

Raqm Numerals are written according to the rules for expressing numbers in Arabic. The largest numeral of a number is written first, except in composite numbers.

The writing of composite numbers (primary numeral + larger numeral) is governed by the following rules:

1. The base form of the primary numeral is used in composite numbers of the same decimal order. The numerals are written transposed, with the primary numeral positioned before the larger numeral. This rule governs the writing of numbers such as 10, 11, ..., 10,000, 11,000, ..., etc.

2. The independent form of the primary numeral is used in composite numbers of different decimal orders. The numerals are written in the regular order with the larger numeral preceding the primary numeral. This rule governs the writing of numbers such as 101, 102, ..., 1,001, 1,002, ..., 1,100, 1,200, ..., etc.

4 The Numerals

4.1 The Primary Unit

The primary unit consists of numerals for the numbers 1 through 9. The numerals are stylized monograms of the Arabic names for the numbers or abbreviations of the names consisting of the initial and one or more letters. The glyphs for Raqm one and two have origins different from those of the numerals in other Siyaq sub-systems, where the numeral one is derived from של מושל יחוד מושל

RAQM	COMPOSITION	ARABIO	C SOURCE	ENGLISH
عم	ع + د + ه	عدد	`dad	one
عيها	ع + د + ن	عددان	ʿdadān	two
یلے	ث + ل + ہے	ثلاثة	<u>t</u> alā <u>t</u> a	three
الم	۱+ر+ع	اربعة	arba ʻa	four
صمه	خ + م	خمسة	<u>h</u> amsa	five
~	س + ہے	ستة	sitta	six
معہ	س + ع + ه	سبعة	sab ʻa	seven
ك	س + ۱ + ہے	ثمانية	<u>t</u> amāniya	eight
لع	ت + ع	تسعة	tis ʿa	nine

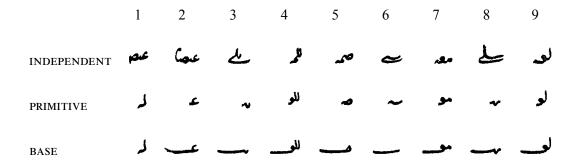
4.1.1 Variant Forms

The following characters have variant forms:

- The numeral are ONE is also written as .
- The numeral Law Two is also written as & and Law.

4.1.2 Base Forms of the Primary Numerals

When written in composite numbers, the primary numbers are written differently, in what might be called their base form. These base forms are derived from the most primitive element of the numeral. This character primitive is joined to a horizontal swash (comparable to _ U+0640 ARABIC TATWEEL) to create the base form. The character primitive is used in the numerals for the hundreds and thousands, with some exceptions as described in the sections for those units.



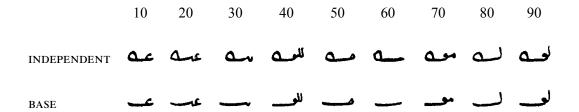
4.2 The Tens Unit

The numerals for the tens unit are composed from the base forms of the primary numerals joined to the tens terminal, which is a stylized form of the $\dot{\upsilon}$ NOON in the Arabic suffix for the tens ($\dot{\upsilon}$) $\bar{u}n$), which is represented as a loop: \bullet . The exceptions are TEN and TWENTY, which have base forms that are used in writing the tens, ten thousands, and millions units. The base form of TEN \bullet and that of TWENTY is \bullet . The base form of TEN \bullet is similar to the base form of TWO (\bullet), which may be distinguished through context.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
عه	عـ + حـد	عشرة	ʿašara	ten
ans	a +	عشرون	ʻišrūn	twenty
سه	a + 	ثلاثون	<u>t</u> alā <u>t</u> ūn	thirty
تلمست	للو_ + ١	اربعون	arba ʿūn	forty
مسه	٠ + 🍑	خمسون	<u>h</u> amsūn	fifty
ه	4 + —	ستون	sittūn	sixty
موه	مو۔۔ + ۵	سبعون	sab ʿūn	seventy
له	۷ + ع	ثمانون	<u>t</u> amānūn	eighty
لحسه	لعب + چه	تسعون	tis ʿūn	ninty

4.2.1 Base Forms of the Tens Numerals

The tens numerals have base forms that are used in the writing of numbers of the ten thousands and millions units. The base form of a tens numeral is derived by dropping the stylized NOON that marks the tens terminal **\(\Delta\)**. Except for TEN, TWENTY, and EIGHTY, the base forms of the tens are identical to the base form for the corresponding primary unit.



4.3 The Hundreds Unit

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
<u></u>	_	مائة	mi ʾa	one hundred
J	_	مائتَان	mi ʾātān	two hundred
لم	(+ ~	ثلاث مائة	<u>t</u> alā <u>t</u> u miʾa	three hundred
۱س)	(_a + ~1	اربع مائة	arbaʻu mi'a	four hundred
صا	(₄ + •	خمس مائة	<u>h</u> amsu mi ʾa	five hundred
4	(+ ~	ستّ مائة	sittu mi 'a	six hundred
ប	(+ J	سبع مائة	sab 'u mi 'a	seven hundred
V	ل + ما	ثمان مائة	<u>t</u> amānu mi ʾa	eight hundred
نعا	لو + ما	تسع مائة	tis 'u mi 'a	nine hundred

The following numerals take forms different from the base shape when used to form the hundred numeral

INDEPENDENT	PRIMITIVE	HUNDREDS		
الم	للو	امر	اس)	400
~	~	~	4	600
معد	مو	1	Ú	700
سلے	~	J	V	800

4.4 The Thousands Unit

The numerals for the thousands are composed from the primitive forms of the primary numerals joined to the terminal —, which is an abbreviation of the Arabic word 'thousand'. When written independently, the thousands are marked as —, where the element represents the absence of other numbers.

It may appear that the thousands are written using the base form of the primary numeral + the element , but this is not the case. Only the TWO THOUSAND follows this pattern: cf. base form of TWO — and TWO THOUSAND —. However, this is on account of the special forms for TWO in each decimal order. Compare base form of SIX — and SIX THOUSAND —.

The forms for one thousand and two thousand have special forms. The one thousand is a monogram of Arabic word 'thousand'; the 'Two thousand is based upon the primitive form of two (2).

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
الستة	_	الف	alf	one thousand
عت	_	الفان	alfān	two thousand
سے	<u>ت</u> + م	ثلاثة الاف	<u>t</u> alā <u>t</u> a ālāf	three thousand
العيت	للو + ست	اربعة الاف	arbaʿa ālāf	four thousand
صت	اند + ح	خمسة الاف	<u>h</u> amsa ālāf	five thousand
حب	ت + ~	ستّة الاف	sitta ālāf	six thousand
موية	مو + یت	سبعة الاف	sabʿa ālāf	seven thousand
س	€ + *	ثمانية الاف	<u>t</u> amāniya ālāf	eight thousand
لع_=	لو + ستّ	تسعة الاف	tis ʿa ālāf	nine thousand

4.4.1 Variant Forms

The following characters have variant forms:

- The numeral الله ONE THOUSAND also takes the shape

4.5 The Ten Thousands Unit

The ten thousands are written using modified forms of the tens numerals, in which the stylized Noon that marks the tens terminal • is dropped. The base forms used in writing the ten thousands unit are derived from modified forms of the tens and are not the base forms of the primary numerals. This assertion is supported by the original Arabic names for these numbers, eg. 80,000 نمائون الفاء <u>tamānūn alfan</u> 'eighty thousands'. Furthermore, the base forms of TEN, TWENTY, and EIGHTY are used in writing these numerals, eg. 80,000 and not **. When written independently, the ten thousands are marked with ", which is a placeholder mark that represents the absence of other numbers.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
مس	ے۔ + ۔۔	عشرة الاف	ʻašara ālāf	ten thousand
عت	/" + 	عشرون الفا	ʻišrūn alfan	twenty thousand
ست	/" + 	ثلاثون الفا	<u>t</u> alā <u>t</u> ūn alfan	thirty thousand
للوس	للو + ""	اربعون الفا	arbaʿūn alfan	forty thousand
وسار	/ * + 	خمسون الفا	<u>h</u> amsūn alfan	fifty thousand
ست	/ " +	ستّون الفا	sittūn alfan	sixty thousand
موس	مو۔۔ + "'	سبعون الفا	sab`ūn alfan	seventy thousand
ات ا	/ " +	ثمانون الفا	<u>t</u> amānūn alfan	eighty thousand
ہوس ک	لعب + ۳	تسعون الفا	tisʿūn alfan	ninty thousand

4.5.1 Variant Forms

The following characters have variant forms:

• The numeral خست TEN THOUSAND also takes the shape خسرة, which is based upon the typology of the thousands. It is derived from the word-initial AIN of عشرة 'ašara' 'ten' + — the thousands base. In some Siyaq sub-systems, TEN is grouped with the primary units and, therefore, its numerals of the various decimal orders are based upon the typology for that group.

4.6 The Hundred Thousands Unit

The numerals for the hundred thousands unit are written as the regular form of the primary unit + HUNDRED THOUSANDS MARK. This method of representing this decimal order in the Raqm tradition differs from the other Siyaq sub-systems in that it borrows from a non-Arabic tradition. The monogram is derived from the Hindi लाख 'hundred thousand'. In modern Indian notation, the *lākh* is written as 1,00,000.

This is different from the Arabic model, where 'hundred thousand' is expressed as مائة الف mi'a alf and various magnitudes of the unit are expressed by prefixing the primary numeral to the unit, eg. 'five hundred thousand' خمس مائة الف 'hamsu mi'a alf.

The Raqm tradition reflects the typology of the Siyaq system in that it has special forms for one hundred thousand and two hundred thousand; these are written as \mathcal{A} and \mathcal{A} , respectively.

RAQM	COMPOSITION	URDU SOURCE		ENGLISH
دیکه	_	ایک لاکھ	ēk lākh	one hundred thousand
UWI	_	دو لا كھ	dō lākh	two hundred thousand
بع مک	ہے + کک	تنين لا كھ	tīn lākh	three hundred thousand
ىل <i>ولىك</i>	الله + کک	چار لاکھ	chār lākh	four hundred thousand
<i>وبک</i>	صمه + کک	پایخ لاکھ	pānch lākh	five hundred thousand
كيسك	ے + کک	چہ لاکھ	chah lākh	six hundred thousand
<i>ىوەلک</i>	معہ + لک	سات لاكھ	sāt lākh	seven hundred thousand
مصامک	لے + لک	آٹھ لاکھ	āťh lākh	eight hundred thousand
<i>توټک</i>	لق + تک	نو لاكھ	nō lākh	nine hundred thousand

The writing of the hundred thousands unit reflects the expression of numbers of the group. The number 300,000 is expressed as and is, therefore, written as (THREE + HUNDRED THOUSANDS MARK), not as * (THREE BASE + HUNDRED THOUSANDS MARK).

4.7 The Millions Unit

The numerals for the millions unit are written with the base form of the tens and HUNDRED THOUSANDS MARK. The millions are an extension of the hundred thousands unit and are expressed as 'tens of hundred-thousands', eg. five million is 'fifty hundred-thousands'. This system is unrelated to the Arabic model, where the millions unit, in the Ottoman Siyaq tradition, is expressed as 'thousand times a thousand', eg. five million unit, in the Ottoman Siyaq tradition, is expressed as 'thousand times a thousand', eg. five million are expressed as 'ten lākh' and written as 10,00,000.

RAQM	COMPOSITION	URDU SOURCE		ENGLISH
ولک	عب + لک	دس لا كھ	das lākh	one million
عراك	Л + — E	بيں لاکھ	bīs lākh	two million
يك	+	تنيس لاكھ	tīs lākh	three million
سوتك	للو + کک	چالس لا كھ	chālis lākh	four million
مک	ب + کک	يجإس لا كھ	pachās lākh	five million
نک	1 + -	ساٹھ لاکھ	sāťh lākh	six million
بولک	مو _ + <i>لک</i>	ستر لاكھ	sattar lākh	seven million
ریک	ل + مک	اسى لا كھ	asī lākh	eight million
يويک	لعب + تک	نية لاكھ	nabbē lākh	nine million

4.8 The Ten Millions Unit

The numbers for the ten millions are an extension of the millions — 'hundreds of hundred-thousands' or 'one-hundred $l\bar{a}kh$ '. In modern Indian notation, the tens millions are expressed as *karoṛ* ('hundred $l\bar{a}kh$ ') and written as 1,00,00,000.

The numeral TEN MILLION is w, which is derived from \sqrt{karor} (Hindi करोड़ karor < Sanskrit कोटि koti). The numeral TWENTY MILLION is written $\sqrt{karoran}$, which is expressed as a plural of $\sqrt{karoran}$ that adds the Persian plural marker \sqrt{a} to $\sqrt{karoran}$. This illustrates an attempt to pattern non-Arabic number names on an Arabic pattern, similar to the method in which 10 and 20 are expressed in Arabic as among airabic $\sqrt{karoran}$ and $\sqrt{karoran}$ and $\sqrt{karoran}$ and $\sqrt{karoran}$.

RAQM	COMPOSITION	URDU	SOURCE	ENGLISH
אנו	_	كرور	karōr	ten million
كودان	_	کروران	karōrān	twenty million

The numbers 30–90 million are written using the primary numeral + the numeral >>> TEN MILLION, eg. 50 million is 'five karor' and is written >>> .

Hypothetically, the method of representing 100 millions would be patterned upon the typology for the ten millions. The number 500 million would be rendered as 'fifty *karor*' and written as ...

4.9 Composite Numbers

Composite numbers of the primary and tens units are written using the base form of the primary numeral and the appropriate tens numeral. The numbers 10–19 are illustrated below. In some Siyaq sub-systems, such

as Diwani, the numbers 10–19 are written using a base form of TEN. Although the Raqm system has a base form of TEN, all composite numbers are written using the regular forms of the tens numerals.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
عه	_	عشرة	ʿašara	ten
دعسه	لم + عبه	احد عشر	aḥad ʿašara	eleven
معه	عد + حد	اثنا عشر	i <u>t</u> nā ʿašara	twelve
مي	صد + س	ثلاثة عشر	<u>t</u> alā <u>t</u> a ʿašara	thirteen
سوعيه	للو_ + عده	اربعة عشر	arbaʿa ʿašara	fourteen
معه	مد + حم	خمسة عشر	<u>h</u> amsa ʿašara	fifteen
عه	- + عه	ستّة عشر	sitta ʻašara	sixteen
سوعی	مو۔ + عبد	سبعة عشر	sab`a `ašara	seventeen
<u>L</u>	مد + حم	ثمانية عشر	<u>t</u> amāniya ʿašara	eighteen
<i>لوع</i> ے	لعب + عيه	تسعة عشر	tisʿa ʿašara	nineteen

When the thousands and ten thousands are written with numerals from smaller decimal orders, the place-holder is dropped and the numerals are stacked: 5,500 \leftarrow FIVE HUNDRED + FIVE HUNDRED + FIFTY THOUSAND.

4.10 Fractions

There are three signs for representing fractions in the Raqm system.

RAQM		ENGLISH
-	1/4	one quarter
•	1/2	one half
-	3/4	three quarters

4.11 Placeholder Mark

4.12 Currency Sign

The character \checkmark is used to write currencies. It represents the rupee currency unit. There is a special orthography for writing currencies in Raqm, which is based on the rupee system that was common across northern South Asia (see Pandey 2007a for a description). This currency system is based upon the rupee $(rupay\bar{a})$, $\bar{a}n\bar{a}$, and $p\bar{a}\bar{\imath}$; with regional variants such as that used in Bengali, which uses the unit $gand\bar{a}$ instead of $p\bar{a}\bar{\imath}$, as shown in Figure 7 (see Pandey 2007b for a description).

When a Raqm numeral is written with the currency mark, it represents rupees: 'fifty rupees' \nearrow . When an Arabic-Indic digit is written with a currency mark, it represents values of the $\bar{a}n\bar{a}$ unit: 'five $\bar{a}n\bar{a}$ ' \nearrow . When a fraction sign is written with the currency mark, the combination represents a fraction of the $\bar{a}n\bar{a}$ unit, which is called the $p\bar{a}\bar{\imath}$ unit: ' $\sqrt{a}n\bar{a}$ ' or 'one $p\bar{a}\bar{\imath}$ ' \nearrow .

5 Implementation

5.1 Encoding Model

Given the above analysis, there are three possible model for encoding the Ragm Numerals.

- 1. Encode each numeral as an atomic character
- 2. Encode the numerals using character primitives
- 3. Encode a combination of numerals and unit marks
- **1. Encode each numeral as an atomic character** The most elementary approach to encoding the Raqm Numerals is to encode each individual numeral as an atomic character. This model would require 85 characters for the numerals: primary units (9), base forms of the primary units (9), tens (9), hundreds (9), thousands (9), ten thousands (9), hundred thousands (9), and millions (9), and ten millions (9); and fractions (3) and currency mark (1).

The advantage of this model is that no special rendering rules are needed to write the numerals. Units larger than millions may be written using combinations of other characters.

The disadvantage is the encoding of redundant characters, in particular the hundred thousands, millions, and ten millions units, which may be written using characters for other units.

2. Encode the numerals using character primitives While the typological characteristics of the Diwani Numerals makes it theoretically possible to encode that Siyaq sub-system using character primitives, this model does not provide an effective means for encoding the Raqm Numerals. In this approach, the Raqm Numerals would be represented using the primitive forms of the primary numerals and the distinctive sign for each decimal order, as is a possibility for Diwani Numerals. However, while there are distinctive signs for tens and other units, there is no such sign for the primary units in Raqm.

To encode Raqm numerals using character primitives, a *PRIMARY UNITS MARK would have to be invented. Raqm one would then be produced by writing *PRIMARY UNITS MARK + J ONE BASE, requiring the rendering engine to produce the correct glyph from the backstore from a given a sequence of characters.

The major disadvantage to this approach is the heavy reliance upon rendering rules. The shaping engine would need to produce the appropriate forms for special ligatures. The number one thousand \longrightarrow would be produced by \longrightarrow THOUSANDS SIGN + \rightarrow ONE BASE.

As the thousands and ten thousands are not distinguished through terminal marks, but by the base shape of the numeral, producing numerals for ten thousands would be expensive in a character-primitives model. The independent form of ten thousand would be produced through the use of multiple terminal marks: PLACEHOLDER MARK + THOUSANDS MARK + TENS MARK + BASE ONE;

Another disadvantage is ordering. With this approach the rendering engine would need to first compose the appropriate number for a base numeral + a unit sign, then order these pairs according to the Arabic counting order.

Although the primitives approach reflects the pattern that underlies the typology of the Raqm Numerals, the complexity of this encoding model will restrict its implementation.

3. Encode a combination of numerals and unit marks A third approach is a mean between the two discussed previously. In this model the numerals of the primary, tens, hundreds, thousands, and ten thousands units are encoded as atomic characters. Based upon their glyphic representation, the various numerals for the hundred thousands and millions may be written using sequences of other characters, eg. the numerals for the hundred thousands unit may be written using the primary unit + ONE HUNDRED THOUSANDS MARK. This model also encodes numerals of various decimal orders, which may be considered typologically unique.

This model would require 64 characters:

- The primary numerals and their combining forms (18)
- The tens (9)
- The hundreds (9)
- The thousands (9)
- The ten thousands (9)
- ONE HUNDRED THOUSAND, TWO HUNDRED THOUSAND, and HUNDRED THOUSANDS MARK (3)
- ONE MILLION and TWO MILLION (2)
- Fraction signs (3)
- Currency marks (1)
- Placeholder mark (1)

Of the three, this approach offers the least complicated method of encoding Ragm Numerals.

5.2 A Basic Character Set for Ragm Numerals

Based upon encoding model #3, 64 characters are required to encode Raqm Numerals in the UCS:

```
xx01 RAQM NUMERAL ONE
xx02 RAOM NUMERAL TWO
xx03 RAQM NUMERAL THREE
xx04 RAQM NUMERAL FOUR
xx05 RAQM NUMERAL FIVE
xx06 RAQM NUMERAL SIX
xx07 RAQM NUMERAL SEVEN
xx08 RAQM NUMERAL EIGHT
xx09 RAQM NUMERAL NINE
xx0A RAQM NUMERAL TEN
xx0B RAOM NUMERAL TWENTY
xx0C RAQM NUMERAL THIRTY
xx0D RAOM NUMERAL FORTY
xx0E RAOM NUMERAL FIFTY
xx0F RAOM NUMERAL SIXTY
xx10 RAQM NUMERAL SEVENTY
xx11 RAQM NUMERAL EIGHTY
xx12 RAOM NUMERAL NINETY
```

```
xx13 RAQM NUMERAL ONE HUNDRED
xx14 RAQM NUMERAL TWO HUNDRED
xx15 RAQM NUMERAL THREE HUNDRED
xx16 RAOM NUMERAL FOUR HUNDRED
xx17 RAQM NUMERAL FIVE HUNDRED
XX18 RAQM NUMERAL SIX HUNDRED
XX19 RAOM NUMERAL SEVEN HUNDRED
xx1A RAQM NUMERAL EIGHT HUNDRED
xx1B RAQM NUMERAL NINE HUNDRED
xx1C RAOM NUMERAL ONE THOUSAND
xx1D RAQM NUMERAL TWO THOUSAND
xx1E RAQM NUMERAL THREE THOUSAND
xx1F RAOM NUMERAL FOUR THOUSAND
xx20 RAQM NUMERAL FIVE THOUSAND
xx21 RAQM NUMERAL SIX THOUSAND
xx22 RAQM NUMERAL SEVEN THOUSAND
xx23 RAOM NUMERAL EIGHT THOUSAND
xx24 RAQM NUMERAL NINE THOUSAND
xx25 RAOM NUMERAL TEN THOUSAND
xx26 RAQM NUMERAL TWENTY THOUSAND
xx27 RAOM NUMERAL THIRTY THOUSAND
xx28 RAQM NUMERAL FORTY THOUSAND
xx29 RAQM NUMERAL FIFTY THOUSAND
xx2A RAOM NUMERAL SIXTY THOUSAND
xx2B RAQM NUMERAL SEVENTY THOUSAND
xx2C RAOM NUMERAL EIGHTY THOUSAND
xx2D RAQM NUMERAL NINETY THOUSAND
xx2E RAQM NUMERAL ONE HUNDRED THOUSAND
xx2F RAQM NUMERAL TWO HUNDRED THOUSAND
xx31 RAQM NUMERAL TEN MILLION
xx32 RAOM NUMERAL TWENTY MILLION
xx33 RAQM NUMERAL COMBINING ONE
xx34 RAQM NUMERAL COMBINING TWO
xx35 RAQM NUMERAL COMBINING THREE
xx36 RAOM NUMERAL COMBINING FOUR
xx37 RAQM NUMERAL COMBINING FIVE
xx38 RAQM NUMERAL COMBINING SIX
xx39 RAOM NUMERAL COMBINING SEVEN
xx3A RAQM NUMERAL COMBINING EIGHT
xx3B RAQM NUMERAL COMBINING NINE
xx3C RAOM FRACTION ONE OUARTER
xx3D RAQM FRACTION ONE HALF
xx3E RAQM FRACTION THREE QUARTERS
xx3F RAOM CURRENCY MARK RUPEE
xx40 RAQM HUNDRED THOUSANDS MARK
xx41 RAQM PLACEHOLDER MARK
```

6 References

- Ahmad Dihlavī, Sayyid. 1974. فر هنگ آصفية [Farhang-i Āsafiyyah]. vol. 2. Delhi: Taraqqī-e-Urdū Board, National Academy.
- Barker, Muhammad Abd-al-Rahman. 1967. *A Course in Urdu*. Vol. 1. Montreal: Institute of Islamic Studies, McGill University.
- Gladwin, Francis. 1790. *A Compendious System of Bengal Revenue Accounts*. In three parts. Part I. Calcutta: Manuel Cantopher.
- Naim, Chaudhry M. 1999. *Introductory Urdu*. 3rd ed. rev. Chicago: South Asia Language & Area Center, University of Chicago.

- Muhazzab, Muhammad Mirza. 195-?. مُحرِّب اللغت [Muhazzab al-Lughat]. Lakhnau, Muhafiz Urdu Book Depot.
- Muqtadirah Qaumi Zaban. 2001. درسي اردو لغت [Darsi Urdu Lughat]. Taba' 1. Silsilah-yi matbu'at-yi Muqtadirah-yi Qaumi Zaban; 391. Islamabad.
- Pandey, Anshuman. 2007a. "Proposal to Encode North Indic Number Forms in ISO/IEC 10646." ISO/IEC JTC1/SC2/WG2 N3367 L2/07-354. October 7, 2007. http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3367.pdf
- ——. 2007b. "Proposal to Encode the Ganda Currency Mark for Bengali in ISO/IEC 10646." ISO/IEC JTC1/SC2/WG2 N3311R L2/07-192R. October 8, 2007. http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3311.pdf
- Platts, John T. 1909. *A Grammar of the Hindūstānī or Urdū Language*. 5th imp. London: C. Lockwood. Stewart, Charles. 1825. *Original Persian Letters and Other Documents with Fac-Similes*. London: printed for the author by W. Nicol.

Table of Rahm.											
رفم	تعسلا	وفسم	تعداد	رفم	نعداد	رفسو	تعداد	رفسم	تعداد	رف	قداد
J.	0	81 16	i i	()		در دی ارالوس	1		40		20
ی د	p	اله		لب		ii .	,	لعده	,		
نها و	F.,	اله		عيد		<u>المع</u> يده		عيد			۲
6		يه		l	l .	لل <u>ي</u> ولا .	1 :	i e	44	l l	۴
l l . I		للو <u>ليه</u> ا		i				للحي			ዮ
٠ 🗘		ب				مل <i>ع</i>			40	صمد	٥
L L		ب	1						44		7
V	1	موليه	۸۷	موي	1	_		موعيه	4 6	معہ	د
لعا	٩	بله		9		للعي				سلے	٨
الستا	١	لوكي	49	لوي	ર્ય ૧	لطحي	P 9	لوعيه	4 4	لعہ	٩
عتا	` v •••	لعسه	٩	مب	د ٠	مـه	۱ . ۵	9	۳.	عه	١.
سمتا	۴	لالحث	٩١	لمعه	د١	لەصت	٥١	ليت	۱ ۳	لعه	• •
العي	۴	ععد	9 1	عي	42	عصيه	۲۵	صد	* 4	ععه	1 4
صمت	٥	سعي	9 4	سعيه	د ۳	عد	۳۵	يحو	۳۳	حي	۱۳
سهي	٩	المولحي	9 6	الموصيه	د ر	للحصيه	م٥	هوسي	م ۳	العصي	۱۴
امعي	٤٠٠.	ملحيه	۹۵	ميده	60	مصه	مه	مي	۳٥	معت	10
سمت	۸	نویه	4 4	مي	۷ ۲	_	0 7	س	٣٦	<u>عـ</u>	١ ٧
لع_=	٩	مولعيه	۹۷	معيه	دد	سوي	ه د	موسي	ے م	پی	۱۷
عمت		پلايه	9 1	اس	د ۸	ہے	۸٥	س	۳۸	جي ا	1 4
لعمست	· · · ·	ايلىيە	49	التعييه	L q	الوصية	٥٩	لوسيه	۳۹	لوعث	١٩
لاكيه	、	6	١	اله	۸.	ه	ч.	للميك	۴.	عده	٧.
,								_			
- to of an ana; j = 2 an ana; j = 2 of an ana; j = one ma											
?ì											
اره برائی از میل مار میل											

Figure 1: Table showing Siyaq forms as used in South Asia (from Platts 1909: 60). It should be noted that the translated values of the Raqm examples are wrong. The value 795 Rs and $11\frac{3}{4}$ As as represented in Raqm is actually 295 Rs and $1\frac{3}{4}$ As. Platts represents \vec{l} as $\frac{1}{2}$ \vec{l} \vec{l} and in the table, but assesses it as $\frac{3}{4}$ \vec{l} \vec{l}

SYMBOL	VALUE	SYMBOL	VALUE	SYMBOL	VALUE
,-	-/-/3	/ •	-/-/9	1	-/1/3
,•	-/-/6	,1	-/1/-	/·1	-/1/6
SYMBOL	VALUE	SYMBOL	VALUE	SYMBOL	VALUE
<u>, :1</u>	-/1/9	معد	12/-/-	معم	70/-/-
۲,	-/2/-	سعيه	13/-/-	لمر	80/-/-
عمر	1/-/-	المعطيب	14/-/-	لعثه	90/-/-
, C	2/-/-	ميە	15/-/-	· , l	100/-
سے	3/-/-	عيه	16/-/-	\lambda	200/-
للعشر	4/-/-	معت	17/-/-	سار	300/-
صمر	5/-/-	رسے ر	18/-/-	للعمار	400/-
, –	6/-/-	لعصم	19/-/-	صمار	500/-
معم	7/-/-	عده	20/-/-	سمار	600/-
, 	8/-/-	, س	30/-/-	معمار	700/-
لعدر	9/-/-	للعث	40/-/-	ر	800/-
عدر	10/-/-	مم	50/-/-	تعمار	900/-
لەعسے	11/-/-	,-	60/-/-	الشمر	1,000/-
				لا كحف	lakh/-

Figure 2: Table showing Siyaq forms as used in South Asia (from Barker 1967: 356–357).

8.6. Sums: Both India and Pakistan now have a decimal coinage system, a rupee being divided into one hundred paisas. In Urdu, the decimal point is wirtten as: 5. Examples:

15 • = Re. 1.00
$$5\Delta$$
• = 50 p. $5 \cdot \Delta$ = 5 p. 1510° = Rs. 1.14

8.7. Before the currency was reformed in the two countries, a rupee was divided into sixteen annas or sixty-four pice (paisa). There was then also a different system, besides the numerals, for writing sums.

Figure 3: Table showing Siyaq forms as used in South Asia (from Naim 1999: 49–50).

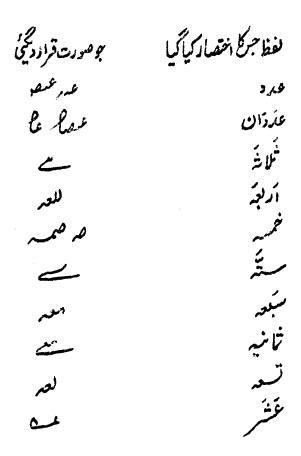


Figure 4: The Arabic sources of the Urdu Siyaq forms (from Muhazzab 195-?: 51).

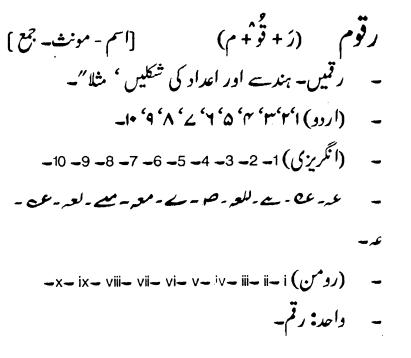


Figure 5: Table showing Siyaq forms as used in South Asia (from Muqtadirah Qaumi Zaban 2001: 718).

Figure 6: Table showing Siyaq forms as used in South Asia (from Dihlavi 1974: 363).

TABLE OF FIGURES.									
Cowriss.	Gundahs.	Gundahs.	Annas.						
* — i	17/16	1/ 1	/1 1						
$\frac{1}{2}$ · 2	17 17	7/ 2	17 2						
³ → 3	1/2 18	°∕ 3	/r 3						
	19/ 19	½ 4	/r 4						
		% 5	/° 5						
		1/ 6	/1 6						
,		77	1 7						
		1 / 8	/^ 8						
		9/9	/9 9						
	ł	1/ 10	/1. 10						
		14 11	/11 11						
		17 12	/17 12						
		117 13	/1º 13						
		15 14	114						
		10/ 15	/1° 15						

Observe, that Annas are distinguished from Gundahs by the stroke being placed to the left of the former, and on the right side of the latter.

Figure 7: Table showing method of writing fractions in South Asian tradition (from Gladwin 1790: 5)

بموالاحد

مولك من من من مولك مولك من من المنا مولك مولك مولك مولك مالك ت به مق رہے رہے ہے سی سی میں میں میں ہی مرس 40,000 30,000 20,000 10,000 9,000 8,000 7,000 6,000 5,000 4,000 500, food 400, food 300, food 200, 000 / 100 food 90,000 80,000 70,000 60,000 000 / 4.000, poo 3,000, foo 2,000 foo 1000, gro 900.000 / 000, foo 700, foo 600, pool مرور مرور مرور کوران العمال لمرکزور کوران

Figure 8: Table showing Siyaq forms as used in South Asia (from Stewart 1825: Plate 7).

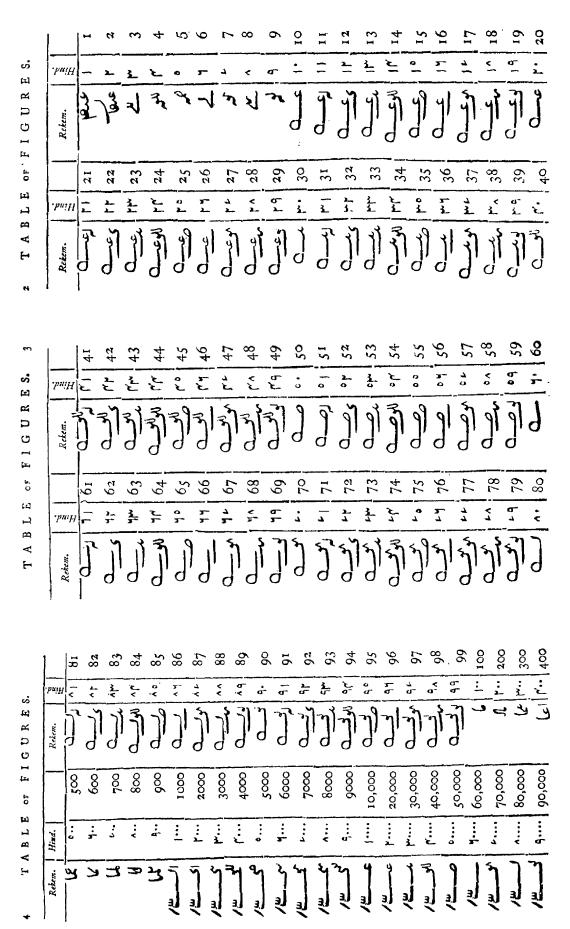


Figure 9: Table showing printed Siyaq forms as used in South Asia (from Gladwin 1790: 2-4).

The Rekem, or Siyak characters, being only contractions of Arabic words, the following Table may ferve to imprefs them on the memory.

Arabic Words.	Rekem.		Arabic Words.	Rokem.		Arabic Words.	R	ckem.	
			<u> </u>		<u> </u>	" " ""	jained.	Separate.	
عثمر	عـــــــــــــــــــــــــــــــــــــ	10	احدعتم	م	11	ء.و	إلم	إعبع	I
عشرين	عـــــ	20	اثنا عشر	عـــد	1 2,	عدوان	کـــ	عنفا	2
ا ثا يبن	یہ	30	أثابثه عشر	مـــد	13	أشر أ	أيب [سع	3
ار بعاین	الوسه	40	ار بعه عشر	الوعيي	14	اربعة	اللوسد ا	أناور	4
نهساین	صــه	50	انمسته عشر	مــــه	15	خمــة	اص	ا همه	5
ستين	ا مـــ	60	ا ستعشر	مت	16	ستة		2	6
سبعين	محت ا	70	اسبعه عشر	موعيده	17	سبعة ا	امو۔ ا	ا کور	7
ثمانين	ره ا	80	انمانية عشرا	م_د ^	18	ثمانيه		سے	8.
تسعين	الحسه	90	[تسعه عشر]	لعيده	19	تسعيم	الحـــ ا	ا لو	9:

N O T E. It is necessary to remark regard-	Arabic Words.	Rekem.		Arabic Words.	Rekem.	
ing the two first digits, that when	الف.	75	1000	مايية	6	100
combined with tens, is a contraction of , and cof (")	الفان	1 <u>m</u>	2000	ماي-نان	N	200
الما الم المسلمة والجرا	ثلاثية آلاف	1 2 1	3000	ثاشمايية	اعا	300
•	اربعم آلاف	/ <u>w</u>	4000	ار بعمايه"	اعا	400
	خمسته آلان	/ <u>w</u> _co	5000	بخب ما يه	.12	500
	سيثة آلاف	1 <u>m</u>	6000	ستعايبة	.K	600
	السبعة آلاف	\ <u>m</u>	7000	سبعمايه	ill	700
	أثمانيه آلاف	/ <u>w</u>	8000	شانرا به	y l	800
	تسعمر آلاف	/ <u>"</u>	9000	تسعمايه	L	900

Figure 10: Table showing the Arabic sources of Siyaq forms (from Gladwin 1790: 6–7).