L2/09-148
2009-04-21

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."[1] A specimen of Raqm printed in Gladwin's metal fonts is given in Figure 9.

[1] Gladwin 1790: vii.

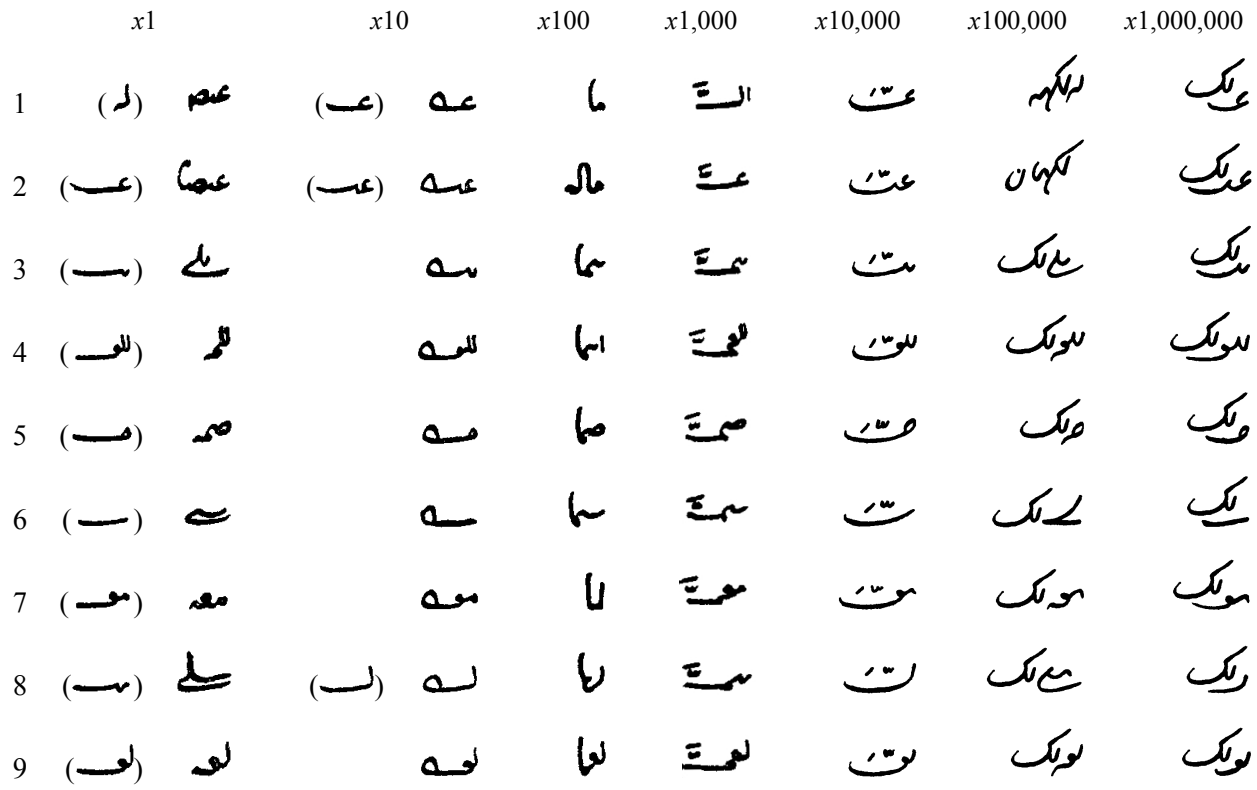


Table 1: South Asian forms of the Siyaq numerals for six decimal orders.

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: ٥ FIVE PRIMITIVE ← ٥ 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: ـ٥ـ FIVE BASE ← ـ BASE MARK + ٥ FIVE PRIMITIVE.
- The numerals for the tens consist of the base form of the primary numeral joined to a distinctive terminal: ٥٠ FIFTY ← ٠ TENS MARKS + ـ٥ـ FIVE BASE.
- The numerals for the hundreds consist of the base form of the primary numeral joined to the numeral for 100: ٥٠٠ FIVE HUNDRED ← ١٠٠ ONE HUNDRED + ٥ 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: ٥٠٠٠ FIVE THOUSAND ← ـ THOUSANDS MARK + ٥ FIVE PRIMITIVE. When written independently, the thousands are marked with the placeholder ٠, eg. ٥٠٠٠ FIVE THOUSAND
- The numerals for the ten thousands consist of a modified (base) form of the tens numerals, which is produced by dropping the ٠ TENS MARKS, and a placeholder when written independently: ٥٠٠٠٠ FIFTY THOUSAND ← ٠ PLACEHOLDER + ـ٥ـ 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 ـ٨ـ and the base form of EIGHT is ـ٨.
- 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: ٥٠٠٠٠٠ FIVE HUNDRED THOUSAND ← ١٠٠٠٠٠ HUNDRED THOUSANDS MARK + ٥ 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: ٥٠٠٠٠٠٠ FIVE MILLION ← ١٠٠٠٠٠ HUNDRED THOUSANDS MARK + ـ٥ـ 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: ٥٠٠٠٠٠٠٠ FIFTY MILLION ← ١٠٠٠٠٠٠٠ TEN MILLION + ٥ 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 احد *aḥad* 'one' and TWO from اثنان *iṯnān* 'two'. The Raqm numeral ONE is based upon Arabic عدد *ʿdad* 'number' and TWO is based upon عددن *ʿdadan* 'dual'.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
	ع + د + ه	عدد	*ʿdad*	one
	ع + د + ن	عددان	*ʿdadān*	two
	ث + ل + ے	ثلاثة	*ṯalāṯa*	three
	ا + ر + ع	اربعة	*arbaʿa*	four
	خ + م	خمسة	*ḫamsa*	five
	س + ے	ستّة	*sitta*	six
	س + ع + ه	سبعة	*sabʿa*	seven
	س + ا + ے	ثمانية	*ṯamāniya*	eight
	ت + ع	تسعة	*tisʿa*	nine

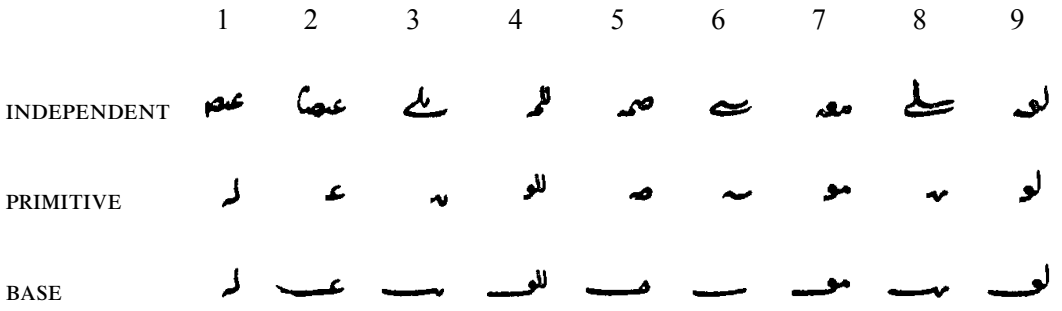
4.1.1 Variant Forms

The following characters have variant forms:

- The numeral  ONE is also written as  .
- The numeral TWO is also written as  and .

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 ن NOON in the Arabic suffix for the tens (اون *ūn*), which is represented as a loop: ـه. 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 is عـ and that of TWENTY is عـ. The base form of TEN عـ is similar to the base form of TWO (عـ), which may be distinguished through context.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
عه	عـ + ـه	عشرة	*ʿašara*	ten
عه	عـ + ـه	عشرون	*ʿišrūn*	twenty
ثه	ثـ + ـه	ثلاثون	*ṯalāṯūn*	thirty
ربه	ربـ + ـه	اربعون	*arbaʿūn*	forty
خمه	خمـ + ـه	خمسون	*ḫamsūn*	fifty
ستّه	ستّـ + ـه	ستّون	*sittūn*	sixty
سبه	سبـ + ـه	سبعون	*sabʿūn*	seventy
ثمه	ثمـ + ـه	ثمانون	*ṯamānūn*	eighty
تسه	تسـ + ـه	تسعون	*tisʿūn*	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

ـه. Except for TEN, TWENTY, and EIGHTY, the base forms of the tens are identical to the base form for the corresponding primary unit.

	10	20	30	40	50	60	70	80	90
INDEPENDENT									
BASE									

4.3 The Hundreds Unit

The numerals for the hundreds are derived from the primitive forms of the primary numerals joined to the mark ـما, which is the abbreviation of the Arabic word مائة 'hundred' as ما. However, not all primitive forms of the primary numerals used in writing the hundreds are identical to those described in section 4.1.2. The primitives used in writing the numerals for FOUR HUNDRED, SIX HUNDRED, SEVEN HUNDRED, and EIGHT HUNDRED represent a different mutation of the primary numerals. These mutations are described below. Apart of these, the other exceptions are the numerals ONE HUNDRED and TWO HUNDRED, which are monograms of the Arabic names for these numbers.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
	—	مائة	*miʾa*	one hundred
	—	مائتَان	*miʾātān*	two hundred
	 +	ثلاث مائة	*ṯalāṯu miʾa*	three hundred
	 +	اربع مائة	*arbaʿu miʾa*	four hundred
	 +	خمس مائة	*ḫamsu miʾa*	five hundred
	 +	ستّ مائة	*sittu miʾa*	six hundred
	 +	سبع مائة	*sabʿu miʾa*	seven hundred
	 +	ثمان مائة	*ṯamānu miʾa*	eight hundred
	 +	تسع مائة	*tisʿu miʾa*	nine hundred

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

		HUNDREDS	PRIMITIVE	INDEPENDENT
400				
600				
700				
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 ـــه, which is an abbreviation of the Arabic word الف 'thousand'. When written independently, the thousands are marked as ـــهـــ, where the element ـــ represents the absence of other numbers.

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

The forms for ONE THOUSAND and TWO THOUSAND have special forms. The ONE THOUSAND is a monogram ـــالـــ of Arabic word الف 'thousand'; the ـــعـــ TWO THOUSAND is based upon the primitive form of TWO (ـــع).

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
ـــالـــ	—	الف	*alf*	one thousand
ـــعـــ	—	الفان	*alfān*	two thousand
ـــسهـــ	ـــه + ـــهـــ	ثلاثة الاف	*ṯalāṯa ālāf*	three thousand
ـــلهـــ	ـــلو + ـــهـــ	اربعة الاف	*arbaʿa ālāf*	four thousand
ـــصهـــ	ـــص + ـــهـــ	خمسة الاف	*ḫamsa ālāf*	five thousand
ـــسهـــ	ـــس + ـــهـــ	ستّة الاف	*sitta ālāf*	six thousand
ـــسعهـــ	ـــسع + ـــهـــ	سبعة الاف	*sabʿa ālāf*	seven thousand
ـــسمهـــ	ـــسم + ـــهـــ	ثمانية الاف	*ṯamāniya ālāf*	eight thousand
ـــلعهـــ	ـــلع + ـــهـــ	تسعة الاف	*tisʿa ālāf*	nine thousand

4.4.1 Variant Forms

The following characters have variant forms:

- The numeral ـــالـــ ONE THOUSAND also takes the shape ـــالله.
- The numeral ـــعـــ TWO THOUSAND also takes the shape ـــاعـــ.

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 ـــن 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 is الفا ثمانون *ṯamānūn alfan* 'eighty thousands'. Furthermore, the base forms of TEN, TWENTY, and EIGHTY are used in writing these numerals, eg. 80,000 ـــسمـــ and not *ـــسنـــ. When written independently, the ten thousands are marked with ـــ, which is a placeholder mark that represents the absence of other numbers.

RAQM	COMPOSITION	ARABIC SOURCE		ENGLISH
𞲕	عـ + [illegible]	عشرة الاف	*ʿašara ālāf*	ten thousand
𞲖	عـ + [illegible]	عشرون الفا	*ʿišrūn alfan*	twenty thousand
𞲗	ـبـ + [illegible]	ثلاثون الفا	*ṯalāṯūn alfan*	thirty thousand
𞲘	للو + [illegible]	اربعون الفا	*arbaʿūn alfan*	forty thousand
𞲙	ھـ + [illegible]	خمسون الفا	*ḫamsūn alfan*	fifty thousand
𞲚	ـ + [illegible]	ستّون الفا	*sittūn alfan*	sixty thousand
𞲛	مو + [illegible]	سبعون الفا	*sabʿūn alfan*	seventy thousand
𞲜	لـ + [illegible]	ثمانون الفا	*ṯamānūn alfan*	eighty thousand
𞲝	لو + [illegible]	تسعون الفا	*tisʿūn alfan*	ninty thousand

4.5.1 Variant Forms

The following characters have variant forms:

- The numeral 𞲕 TEN THOUSAND also takes the shape عـ, which is based upon the typology of the thousands. It is derived from the word-initial AIN of عشرة *ʿašara* 'ten' + ـ 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 + لک 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 لک 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 *miʾa alf* and various magnitudes of the unit are expressed by prefixing the primary numeral to the unit, eg. 'five hundred thousand' خمس مائة الف *ḫamsu miʾa 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 لکھ and لکھن , respectively.

RAQM	COMPOSITION	URDU SOURCE		ENGLISH
	—	ایك لاكھ	*ēk lākh*	one hundred thousand
	—	دو لاكھ	*dō lākh*	two hundred thousand
	سے + لک	تین لاكھ	*tīn lākh*	three hundred thousand
	چہ + لک	چار لاكھ	*chār lākh*	four hundred thousand
	صد + لک	پانچ لاكھ	*pānch lākh*	five hundred thousand
	چہ + لک	چہ لاكھ	*chah lākh*	six hundred thousand
	ہفت + لک	سات لاكھ	*sāt lākh*	seven hundred thousand
	ہشت + لک	آٹھ لاكھ	*āṫh lākh*	eight hundred thousand
	نہ + لک	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 تین لاكھ, 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 خمسة الاف مرة الف *ẖamsu ālāf maratan alf* 'five-thousand times a thousand'. In modern Indian notation, the millions are expressed as 'ten *lākh*' and written as 10,00,000.

RAQM	COMPOSITION	URDU SOURCE		ENGLISH
	ع + لک	دس لاکھ	*das lākh*	one million
	عع + لک	بیس لاکھ	*bīs lākh*	two million
	ت + لک	تیس لاکھ	*tīs lākh*	three million
	الف + لک	چالس لاکھ	*chālis lākh*	four million
	پ + لک	پچاس لاکھ	*pachās lākh*	five million
	س + لک	ساٹھ لاکھ	*sāṫh lākh*	six million
	ست + لک	ستّر لاکھ	*sattar lākh*	seven million
	ا + لک	اسی لاکھ	*asī lākh*	eight million
	ن + لک	نبّے لاکھ	*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 *lākh*'. In modern Indian notation, the tens millions are expressed as *karoṛ* ('hundred *lākh*') and written as 1,00,00,000.

The numeral TEN MILLION is  , which is derived from کرور *karor* (Hindi करोड़ *karoṛ* < Sanskrit कोटि *koṭi*). The numeral TWENTY MILLION is written  *karorān*, which is expressed as a plural of that adds the Persian plural marker آن *ān* to کرور. 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 عشرة *ʿašara* and عشرون *ʿišrūn*.

RAQM	COMPOSITION	URDU SOURCE		ENGLISH
	—	کرور	*karōr*	ten million
	—	کروران	*karōrān*	twenty million

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

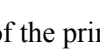
Hypothetically, the method of representing 100 millions would be patterned upon the typology for the ten millions. The number 500 million would be rendered as 'fifty *karoṛ*' and written as .

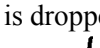
4.9 Composite Numbers

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
	—	عشرة	*ʿašara*	ten
	 +	احد عشر	*aḥad ʿašara*	eleven
	 +	اثنا عشر	*iṯnā ʿašara*	twelve
	 +	ثلاثة عشر	*ṯalāṯa ʿašara*	thirteen
	 +	اربعة عشر	*arbaʿa ʿašara*	fourteen
	 +	خمسة عشر	*ḫamsa ʿašara*	fifteen
	 +	ستّة عشر	*sitta ʿašara*	sixteen
	 +	سبعة عشر	*sabʿa ʿašara*	seventeen
	 +	ثمانية عشر	*ṯamāniya ʿašara*	eighteen
	 +	تسعة عشر	*tisʿa ʿašara*	nineteen

As the table above illustrates, the base form of the primary numeral is positioned in a particular fashion with the larger numeral. If the base form of the primary numeral has no vertical element on the right edge, it is written beneath the larger numeral:  16 ←  TEN +  SIX BASE. If a vertical element is present, the swash of the primary numeral subtends beneath the numeral: 19 ← TEN + NINE BASE.

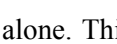
When the thousands and ten thousands are written with numerals from smaller decimal orders, the placeholder is dropped and the numerals are stacked:  5,500 ←  FIVE HUNDRED +  FIVE THOUSAND;  50,500 ← FIVE HUNDRED +  FIFTY THOUSAND.

4.10 Fractions

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

RAQM		ENGLISH
	¼	one quarter
	½	one half
	¾	three quarters

4.11 Placeholder Mark

As described in sections 4.4 and 4.5, the thousands and ten thousands are written when the placeholder mark when they appear alone. This mark is also written as  and  . This mark represents the absence of numbers and is written above the horizontal stroke of numerals:  ← + .

4.12 Currency Sign

The character ⁄ 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 (*rupayā*), *ānā*, and *pāī*; with regional variants such as that used in Bengali, which uses the unit *gaṇḍā* instead of *pāī*, 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' ⁄٥٠. When an Arabic-Indic digit is written with a currency mark, it represents values of the *ānā* unit: 'five *ānā*' ⁄٥. When a fraction sign is written with the currency mark, the combination represents a fraction of the *ānā* unit, which is called the *pāī* unit: '¼ *ānā*' or 'one *pāī*' ⁄ ‾.

5 Implementation

5.1 Encoding Model

Given the above analysis, there are three possible model for encoding the Raqm 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 milllions (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 ١ would then be produced by writing *PRIMARY UNITS MARK + ١ 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 ١٠٠٠ would be

produced by ـــ THOUSANDS SIGN + ـــ 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 ـــ would be produced through the use of multiple terminal marks: ـــ PLACEHOLDER MARK + ـــ THOUSANDS MARK + ـــ TENS MARK + ـــ 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 Raqm Numerals.

5.2 A Basic Character Set for Raqm Numerals

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

```
xx01 RAQM NUMERAL ONE
xx02 RAQM 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 RAQM NUMERAL TWENTY
xx0C RAQM NUMERAL THIRTY
xx0D RAQM NUMERAL FORTY
xx0E RAQM NUMERAL FIFTY
xx0F RAQM NUMERAL SIXTY
```

```
xx10 RAQM NUMERAL SEVENTY
xx11 RAQM NUMERAL EIGHTY
xx12 RAQM NUMERAL NINETY

xx13 RAQM NUMERAL ONE HUNDRED
xx14 RAQM NUMERAL TWO HUNDRED
xx15 RAQM NUMERAL THREE HUNDRED
xx16 RAQM NUMERAL FOUR HUNDRED
xx17 RAQM NUMERAL FIVE HUNDRED
XX18 RAQM NUMERAL SIX HUNDRED
XX19 RAQM NUMERAL SEVEN HUNDRED
xx1A RAQM NUMERAL EIGHT HUNDRED
xx1B RAQM NUMERAL NINE HUNDRED

xx1C RAQM NUMERAL ONE THOUSAND
xx1D RAQM NUMERAL TWO THOUSAND
xx1E RAQM NUMERAL THREE THOUSAND
xx1F RAQM NUMERAL FOUR THOUSAND
xx20 RAQM NUMERAL FIVE THOUSAND
xx21 RAQM NUMERAL SIX THOUSAND
xx22 RAQM NUMERAL SEVEN THOUSAND
xx23 RAQM NUMERAL EIGHT THOUSAND
xx24 RAQM NUMERAL NINE THOUSAND

xx25 RAQM NUMERAL TEN THOUSAND
xx26 RAQM NUMERAL TWENTY THOUSAND
xx27 RAQM NUMERAL THIRTY THOUSAND
xx28 RAQM NUMERAL FORTY THOUSAND
xx29 RAQM NUMERAL FIFTY THOUSAND
xx2A RAQM NUMERAL SIXTY THOUSAND
xx2B RAQM NUMERAL SEVENTY THOUSAND
xx2C RAQM 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 RAQM NUMERAL TWENTY MILLION

xx33 RAQM NUMERAL COMBINING ONE
xx34 RAQM NUMERAL COMBINING TWO
xx35 RAQM NUMERAL COMBINING THREE
xx36 RAQM NUMERAL COMBINING FOUR
xx37 RAQM NUMERAL COMBINING FIVE
xx38 RAQM NUMERAL COMBINING SIX
xx39 RAQM NUMERAL COMBINING SEVEN
xx3A RAQM NUMERAL COMBINING EIGHT
xx3B RAQM NUMERAL COMBINING NINE

xx3C RAQM FRACTION ONE QUARTER
xx3D RAQM FRACTION ONE HALF
xx3E RAQM FRACTION THREE QUARTERS

xx3F RAQM CURRENCY MARK RUPEE

xx40 RAQM HUNDRED THOUSANDS MARK
xx41 RAQM PLACEHOLDER MARK
```

6 References

Ahmad Dihlavī, Sayyid. 1974. فرہنگ آصفیۃ [*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-?. محزّب اللغت [*Muhazzab al-Lughat*]. Lakhnau, Muhafiz Urdu Book Depot.

Muqtadirah Qaumi Zaban. 2001. درسی 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.

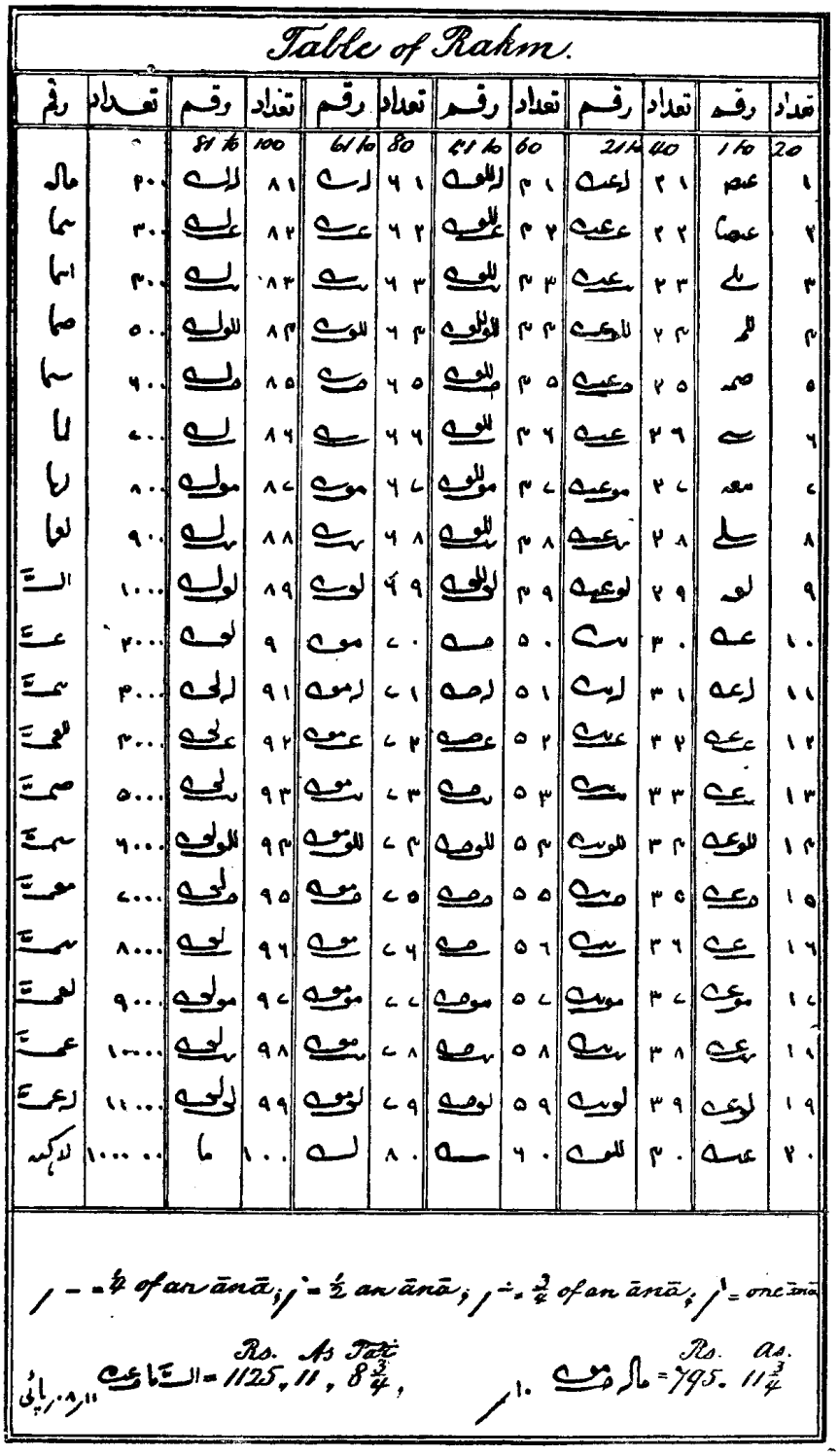


Figure 1: Table showing Siyaq forms as used in South Asia (from Platts 1909: 60). It should be noted that the translated values of the Raqm examples are wrong. The value 795 *Rs* and 11¾ *As* as represented in Raqm is actually 295 *Rs* and 1¾ *As*. Platts represents ⁄ as ½ *ānā* in the table, but assesses it as ¾ *ānā* in the examples.

SYMBOL	VALUE	SYMBOL	VALUE	SYMBOL	VALUE
ـر	-/-/3	شر	-/-/9	١ر	-/1/3
ـ٠ر	-/-/6	ار	-/1/-	١٠ر	-/1/6

SYMBOL	VALUE	SYMBOL	VALUE	SYMBOL	VALUE
١٠ر	-/1/9	عشر	12/-/-	معشر	70/-/-
٢ر	-/2/-	بعشر	13/-/-	لشر	80/-/-
عشر	1/-/-	للعشر	14/-/-	لعشر	90/-/-
غار	2/-/-	مشر	15/-/-	مار	100/-
سے ر	3/-/-	عشر	16/-/-	مار	200/-
للعر	4/-/-	معشر	17/-/-	سار	300/-
صر	5/-/-	مشر	18/-/-	للعمار	400/-
سے ر	6/-/-	لعشر	19/-/-	صمار	500/-
معر	7/-/-	عشر	20/-/-	سمار	600/-
سے ر	8/-/-	سشر	30/-/-	معمار	700/-
لعر	9/-/-	للعشر	40/-/-	لسار	800/-
عشر	10/-/-	صشر	50/-/-	لعمار	900/-
لہ عشر	11/-/-	سشر	60/-/-	الله ر	1,000/-
				لاکھ	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: ٫ .Examples:

١٫٠ = Re. 1.00 ٫٥٠ = 50 p. ٫٠٥ = 5 p. ١٫١٤ = Rs. 1.14

8.7. Before the currency was refomed 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.

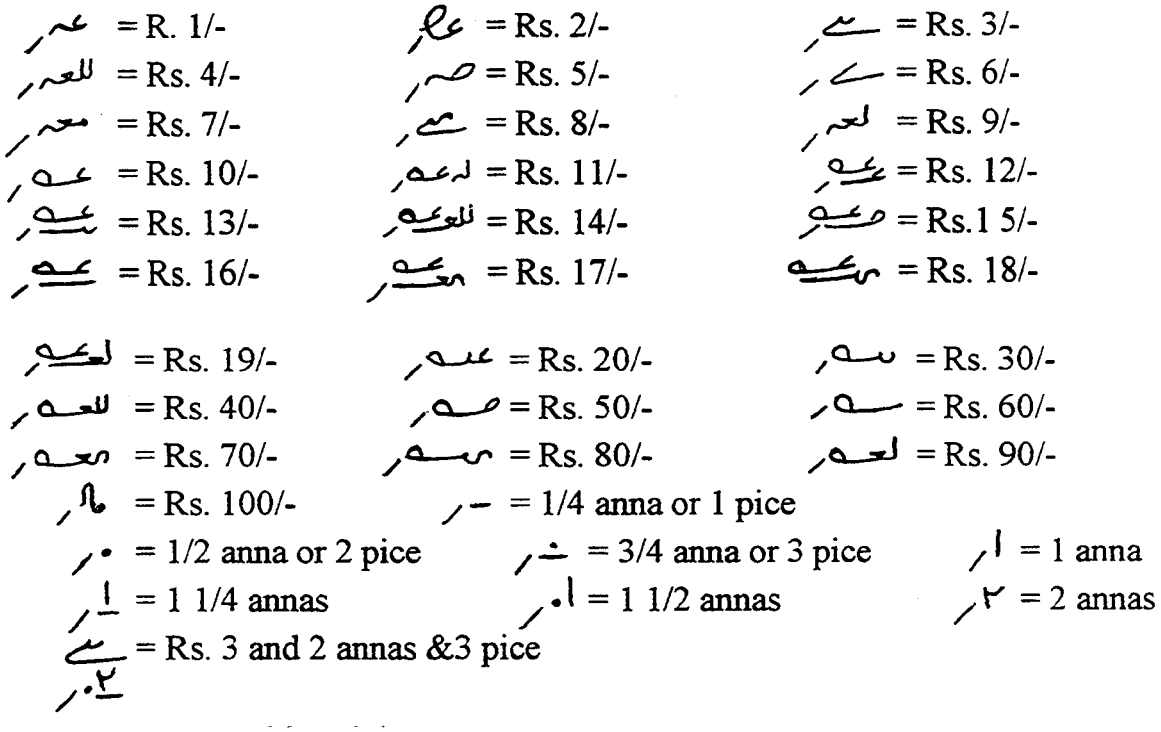


Figure 3: Table showing Siyaq forms as used in South Asia (from Naim 1999: 49–50).

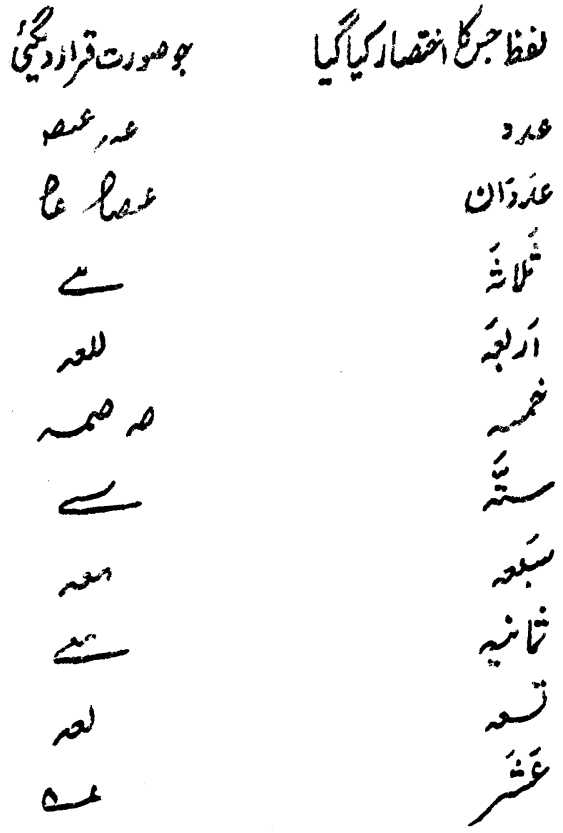


Figure 4: The Arabic sources of the Urdu Siyaq forms (from Muhazzab 195-?: 51).

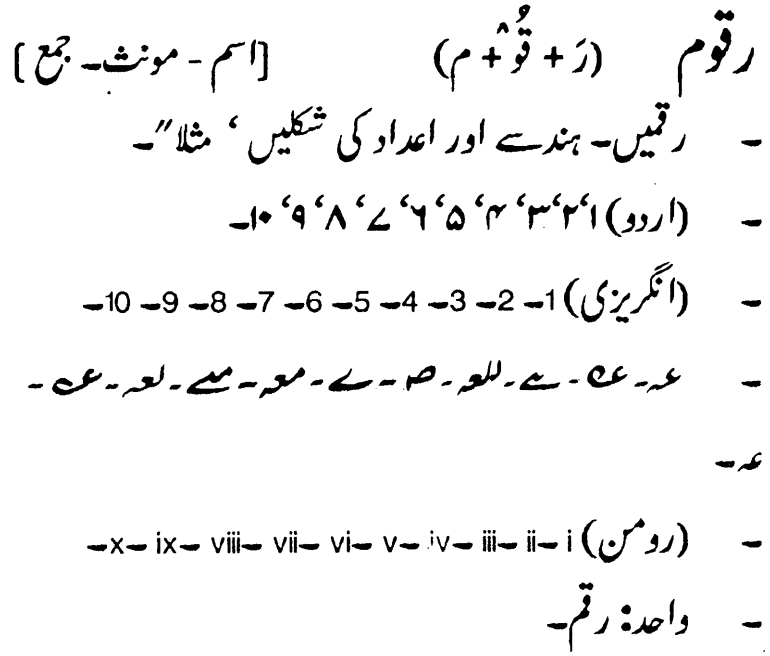


Figure 5: Table showing Siyaq forms as used in South Asia (from Muqtadirah Qaumi Zaban 2001: 718).

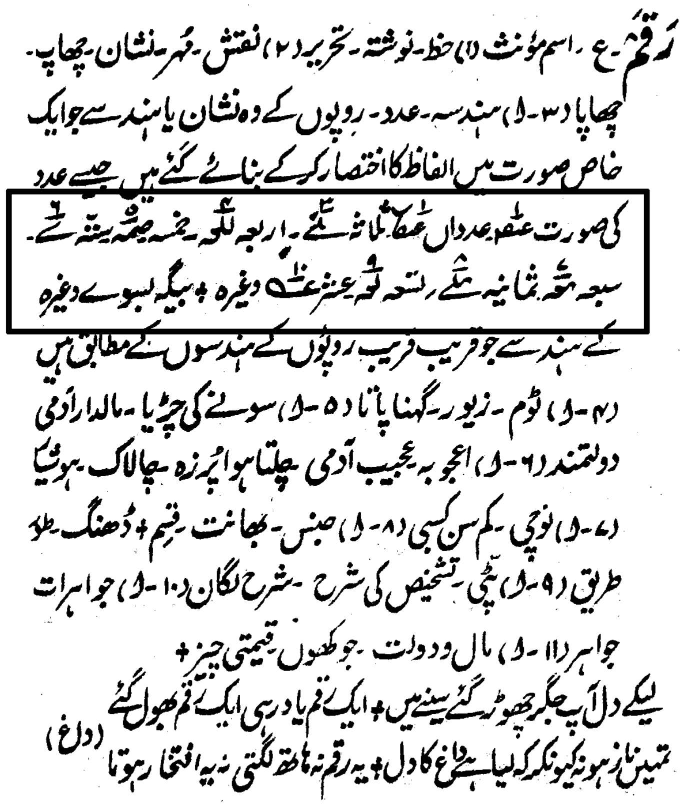

رَقم - ع - اسم مؤنث (۱) خط - نوشتہ - تحریر (۲) نقش - مُہر - نشان - چھاپ -
چھاپا (۳-۱) ہندسہ - عدد - روپوں کے وہ نشان یا ہندسے جو ایک
خاص صورت میں الفاظ کا اختصار کرکے بنائے گئے ہیں جیسے عدد
کی صورت ... اربعہ ... خمسہ ... ستہ ...
سبعہ ... ثمانیہ ... تسعہ ... عشر ... وغیرہ + بیگہ بسوے وغیرہ
کے ہندسے جو قریب قریب روپوں کے ہندسوں کے مطابق ہیں
(۴-۱) ڈوم - زیور - گہنا پاتا (۵-۱) سونے کی چڑیا - مالدار آدمی
دولتمند (۶-۱) اعجوبہ - عجیب آدمی چلتا ہوا پرزہ - چالاک - ہوشیار
(۷-۱) نوچی - کم سن کسبی (۸-۱) جنس - بھانت - قسم + ڈھنگ - طور
طریق (۹-۱) پٹی - تشخیص کی شرح - شرح لگان (۱۰-۱) جواہرات
جواہر (۱۱-۱) مال و دولت - جو کھوں قیمتی چیز +
لیکے دل آپ جگر چھوڑ گئے سینے میں + ایک رقم یاد رہی ایک رقم بھول گئے
تمہیں ناز ہو نہ کیونکر کہ لیا ہے داغ کا دل + یہ رقم نہ ہاتھ لگتی نہ یہ افتخار ہوتا (داغ)

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

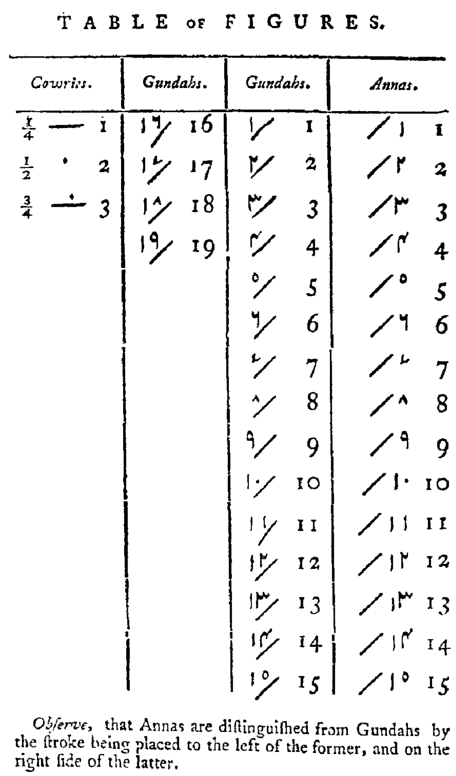


TABLE OF FIGURES. 5

Cowries.		Gundahs.		Gundahs.		Annas.	
¼	1	١٦/	16	١/	1	/١	1
½	2	١٧/	17	٢/	2	/٢	2
¾	3	١٨/	18	٣/	3	/٣	3
		١٩/	19	٤/	4	/٤	4
				٥/	5	/٥	5
				٦/	6	/٦	6
				٧/	7	/٧	7
				٨/	8	/٨	8
				٩/	9	/٩	9
				١٠/	10	/١٠	10
				١١/	11	/١١	11
				١٢/	12	/١٢	12
				١٣/	13	/١٣	13
				١٤/	14	/١٤	14
				١٥/	15	/١٥	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)

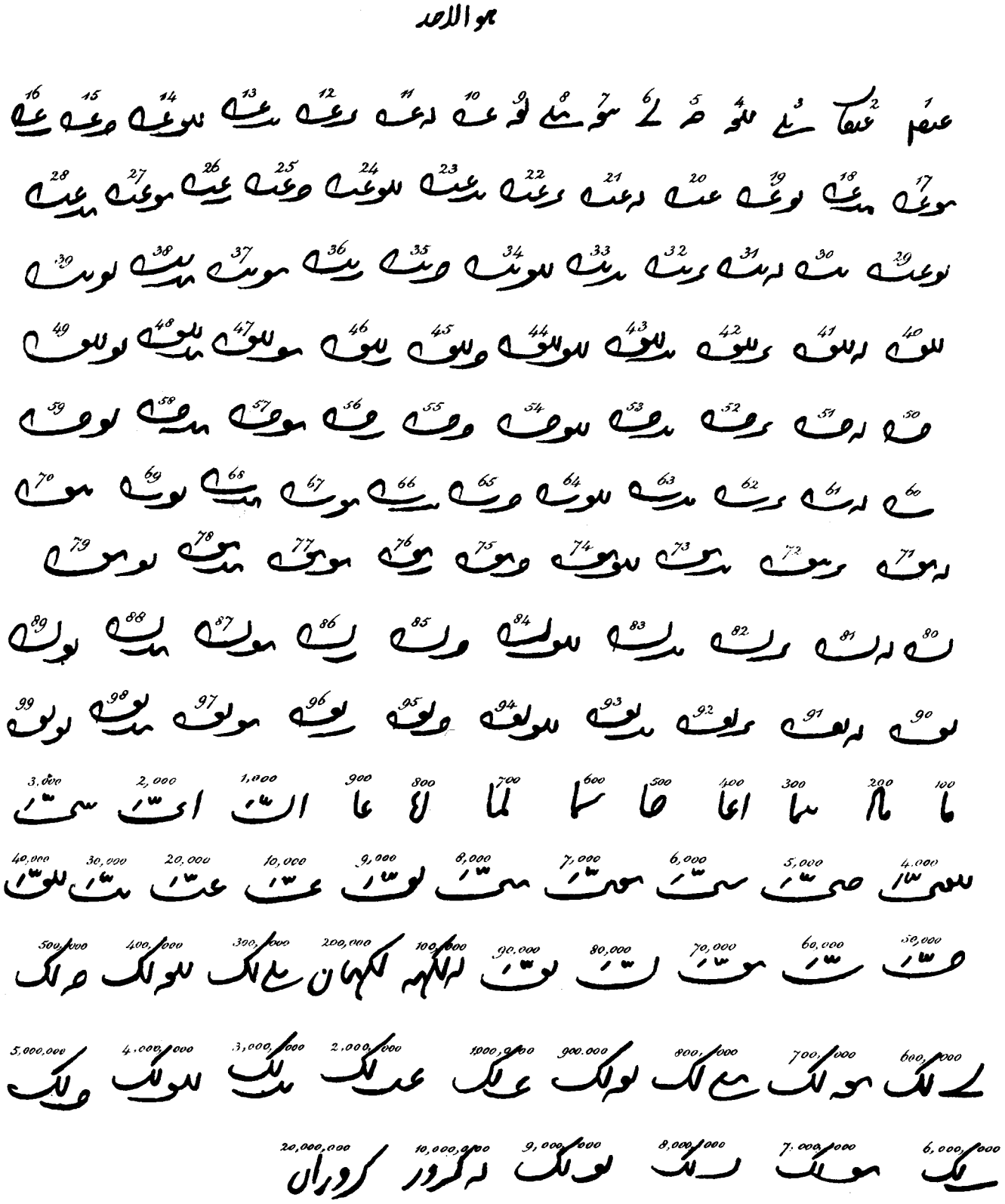


Figure 8: Table showing Siyaq forms as used in South Asia (from Stewart 1825: Plate 7).

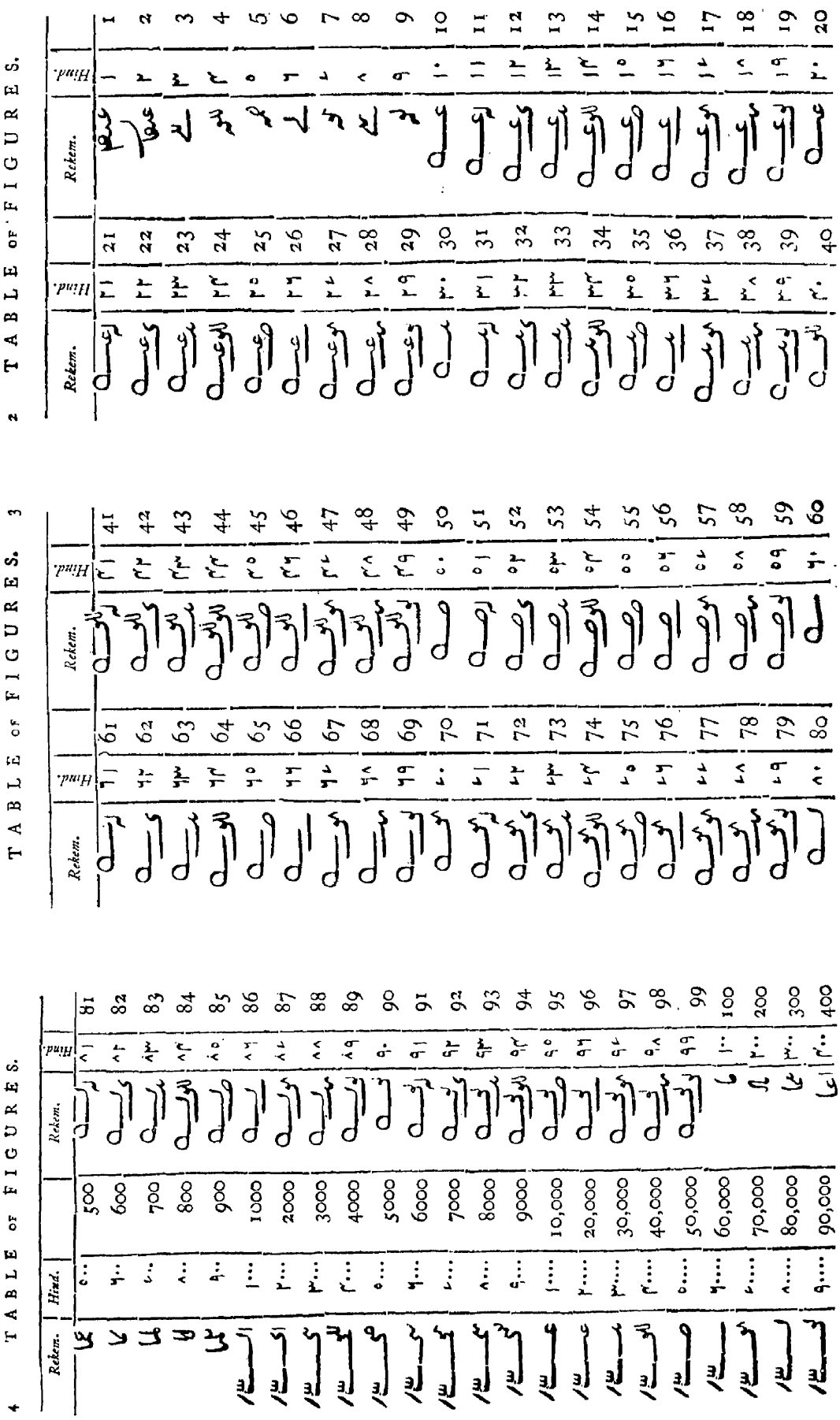


Figure 9: Table showing printed Siyaq forms as used in South Asia (from Gladwin 1790: 2–4).

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

Arabic Words.	Rekem.		Arabic Words.	Rekem.		Arabic Words.	Rekem. joined.	Rekem. ſeparate.	
عشر	[illegible]	10	احد عشر	[illegible]	11	[illegible]	[illegible]	[illegible]	1
عشرين	[illegible]	20	اثنا عشر	[illegible]	12	[illegible]	[illegible]	[illegible]	2
ثلثين	[illegible]	30	ثلثة عشر	[illegible]	13	ثلثة	[illegible]	[illegible]	3
اربعين	[illegible]	40	اربعة عشر	[illegible]	14	اربعة	[illegible]	[illegible]	4
خمسين	[illegible]	50	خمسة عشر	[illegible]	15	خمسة	[illegible]	[illegible]	5
ستين	[illegible]	60	ستة عشر	[illegible]	16	ستة	[illegible]	[illegible]	6
سبعين	[illegible]	70	سبعة عشر	[illegible]	17	سبعة	[illegible]	[illegible]	7
ثمانين	[illegible]	80	ثمانية عشر	[illegible]	18	ثمانية	[illegible]	[illegible]	8
تسعين	[illegible]	90	تسعة عشر	[illegible]	19	تسعة	[illegible]	[illegible]	9

N O T E.

It is neceſſary to remark regarding the two firſt digits, that when combined with tens, [illegible] is a contraction of احد, and [illegible] of اثنا

Arabic Words.	Rekem.		Arabic Words.	Rekem.	
الف	[illegible]	1000	مائة	[illegible]	100
الفان	[illegible]	2000	مائتان	[illegible]	200
ثلثة آلاف	[illegible]	3000	ثلثمائة	[illegible]	300
اربعة آلاف	[illegible]	4000	اربعمائة	[illegible]	400
خمسة آلاف	[illegible]	5000	خمسمائة	[illegible]	500
ستة آلاف	[illegible]	6000	ستمائة	[illegible]	600
سبعة آلاف	[illegible]	7000	سبعمائة	[illegible]	700
ثمانية آلاف	[illegible]	8000	ثمانمائة	[illegible]	800
تسعة آلاف	[illegible]	9000	تسعمائة	[illegible]	900

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