Deciphering Indus Script

Kaleemullah Lashari

The Rosetta Stone a tri-script inscription found in the desert of Egypt, near the village Rashid, by French soldiers, has acquired a meaning that denotes the sort of key, a decoding mechanism, a formula that may assist solve the riddle. But as you know even the Rosetta Stone with the three parallel scripts and two languages incised on it, didn't cause an instant decipherment of Hieroglyphs. It took around 60 years to solve the difficulties involved in the reading of the Egyptian story. Anyone considering the decipherment of Indus script/ signs/pictographs etc. in a century's time may very well be considered over enthusiast.

The enigmatic Indus Script has remained a challenge since quite sometimes, and it is likely that it may evade the attempts of the archaeologists, linguists and scientists yet for another considerable period.

Since 1872-3 when the first so called Indus Seal was found from Harappa it was instantly felt that it represented a new script, which was associated with an ancient language.

So for many attempts have been made to its decipherment, bringing back very interesting set of argument. The reasoning based on the comparative study with the ancient scripts, the efforts to identify the language it represents was the major focus.

It was argued that it holds the key to under stand the great Indian civilization; the comparison was made with many ancient scripts, which included the one nearer home like Brahmi, and such examples situated as far away as Easter Island, and also the Old Slavic; the list is not only long but pretty interesting. Dozens and dozens of the acclaimed decipherments were brushed aside by ever emerging scholars, but there has never been any dearth of new attempts at it.

Impression of a square Indus seal from Harappa (H-103) inscribed along three sides.

It was never doubted that the Indus was a literate civilization (Kenoyer 1998; Possehl 2002), no one seems to doubt this fact that it was not representing language/languages (Daniels and Bright 1996; Pope 1999; Robinson 2002).

Impression of a seal (M-735) from Mohenjodaro with comparing of the last inscribed sign.

The hope for finding larger inscriptions, say a library of Indus, is a dream desire of many of our colleagues, or else finding any multi lingual/script piece may help take us on a new course of discoveries.

Impression of a longest continuous Indus inscription a seal from Mohenjodaro (M-314)

As we know very well that the scholars based in the sub-continent as well as abroad remained involved in such efforts from very beginning.

Fairservis's model of decipherment based on the Dravidian hypothesis (published in 1992) has not been taken seriously because of his lack of familiarity with the Dravidian languages and linguistic techniques.

Inscription consisting of just the sign "7"+ fish on an Indus seal from Harappa (H-9)

There is no doubt that the so called Dravidian model remained a very convincing way, till the end of the century, to be followed for many more attempts.

Some very serious efforts went on to point out that the Indus Script represents Dravidian language/languages.

The computer arrived on the scene in the mid-Sixties. Its contribution in computing the positions of Indus signs was greatly sought. The computer-based analyses of regularities observed in the sign positions were considered convincing and it was believed that these signs represented a spoken language.

Many embarked on this journey; these included a prominent Soviet scholar whose work has also remained associated with decipherment of Mayan writing (Knorozov 1965; 1968) Knorozov published a series of papers entitled Protoindica, and a team of Finnish linguists and computation experts (Parpola, Koskenniemi, Parpola, Aalto 1969) embarked on a project that protracted for a considerable period.

The Parpola and the team went on to state that they had 'broken the code' (Parpola 1970).

Impression of a seal from Mohenjodaro (M-66) with cramping of the last sign of the line and one sign written beneath them.

Despite differences in detail there is a clear convergence of results flowing from the Soviet, Finnish and Indian computer-aided structural analyses.

The Finnish team also made use of computational linguistic techniques to deal with structural problems like word-division procedures and synctical analysis. There is already well-established evidence available proving the general direction of the Indus script to be from right to left.

This evidence is provided by the shorter inscriptions starting at the right edge leaving blank space nearer the left edge, and the displacement of the left-most signs of the longer inscriptions to the second line for want of space. Internal evidence for the direction of writing is obtained by comparing single-line and two-line sequences of identical inscriptions.

The history of efforts at its decipherment is as rich as it is interesting; many claims of its decoding has been very ably reviewed by scholars, more elaborately by Gregory L Possehl (see also [Mahadevan, 2002; Parpola, 2005]). It is a fact that all the 'decipherments' remained unconvincing with no universal consensus on any of the interpretations.

We here were much exposed to one of the remarkable attempts in the recent past, the one made by Asko Parpola. The effort was well supported both by India and Pakistan. But there has not been substantial break through; it in a way sounded like a failure, eating away the hopes of any success in the years to come.

Experts considered implausible identification of pictorial signs, arbitrary assignment of values to non-pictorial signs and diacritic-like marks unconvincing. Even the classification of composite and variant signs was considered to be not beyond doubt. [Harappan.com]

Parpola was considered to be slightly obsessed with the ideas of the 'Harappan religion', and the inexplicable absence of matters relating to the social life and administration of the Harappan polity was considered as weak part in his scheme of decipherment.

It was clear that his interpretations rely more on mythology than on textual or linguistic analysis. There has been another school of thought, which expounded theory of the Indo-Aryan hypothesis. Parpola took no notice

of these models, presumably because there is hardly anything in common between them and his own work. Certain class of scholars described his work as more or less a treatise on Harappan religion rather than decipherment of the Indus script. The absence of any convincing decipherment raised many skeptical comments; the argument was rather strong and it went up to the level to even deny the litracy in the Indus Valley in the third millennium BCE.

The absence of the longer text was at the base of it, this criticism even lost sight of the fact that there were pretty established centralized administrative structures in the Harappan cities and these can not possibly be made functional with out any formal communicational link.
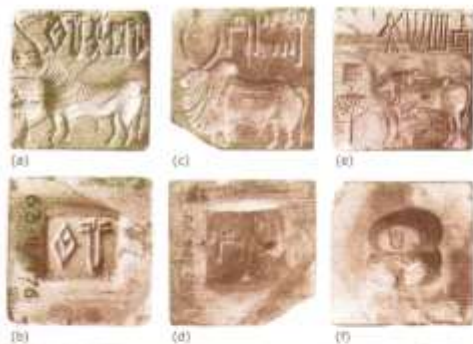
Three seals with an inscription on the knob of the reverse repeating (in the case of f only partly) the initial portion of the inscription on the obvesre (a-b) H-102 from Harappa; (c-d) M-318; (e-f) M-1203 from Mohenjodaro.

Granting that the seal-texts are probably only strings of names and titles, and assuming that the writing is mostly logo-graphic, it would still be necessary to employ minimally parts of speech like pronouns, conjunctions and verbal participles and also grammatical morphs to indicate person, number, gender and case [Harappan.com].

There has been scholars who have taken unconventional course and have made very interesting attempts at understanding the character of the Indus signs.

The controversy over the status of the Indus signs has been quite grim, line of arguments, outlining the reasons for failure of its being understood are interesting as well as outlining many controversies. For example a group of scholars considers it to be the signs only, not representing any language.
The other group considers that non-decipherment is simply because there is no bilingual inscription available. They also argue that the absence of long enough writing is minimizing the chances of its decipherment.

In actuality no knowledge exists about the kings, or the names of the cities, thus looking for such names may help devising some key to understand the characters.

The objects on which these signs appear are 'in terms of art, aesthetic sense and expressions of symmetric, geometric as well as abstract patterns, these objects are unsurpassed in their quality [Yadav & Vahia, 2011; Vahia & Yadav, 2010];

Major among these are the seals, 'inscribed with the Indus script (these) are generally a few square centimeters in size. They are catalogued in the three volumes of the Corpus of Indus Seals and Inscriptions [Joshi & Parpola, 1987; Shah & Parpola, 1991; Parpola, et al., 2010].

Some examples of Indus seals with the Indus script (Copyright Harappa Archaeological Research Project/J.M. Kenoyer, Harappa.com, Courtesy Dept. of Archaeology and Museums, Govt. of Pakistan).

The matter of great interest is that even the usage and purpose of the material on which the inscriptions are appearing is not clear [Possehl, 1996, 20]. The Indus signs are present on variety of artifacts like Seals, sealings, tablets, pottery etc. Seals, as these are called, due mainly to its having the hub like device at its back have been unearthed from number of sites, larger quantity comes from Mohenjodaro. The script therefore holds a vital clue to understanding the Indus culture.

Of all the attempts many were clearly very systematic, logical and were argued well; few were such that could be termed below the scholarly tradition. However all these could not agree to any common point nor could help resolve the issues of meaning and purpose.

'Reasons that make the problem of Indus script challenging are the brevity of the Indus texts, paucity of the data, lack of definitive knowledge about the language(s) that the Indus people spoke, and absence of bilingual inscriptions.... "[Yadav]" It is the reason that many people consider that the Indus script may remain un deciphered for another score of years. Yadev and team worked to understand the basic construction and mechanics of Indus sign. They are of the view that 'the empirical frameworks are unlikely to provide satisfactory decipherment without a proper understanding of the syntax to guide and validate such frameworks in an objective manner. They 'used various computational techniques to identify aggregate characteristics of the Indus script without making any assumptions about its content' [Yadav].

They employed computational techniques related to machine learning, data minning and information theory to understand the characterstics and the syntax of Indus script; The team thinks this '..will provide an objective testing ground for any claims of decipherment of Indus script [Yadav].
They argued that systematic writing would have specific ordering of signs. The frequency of signs and sign combinations as well as their pattern of appearance in the texts elucidate the syntax of the writing.

Analysis of syntactic patterns

The study went on to suggest 'limited number of signs for text beginners and enders and significant constraints on the pairing of signs.

'This suggests that in the Indus writing only a small number of signs were allowed to end the texts while a relatively larger number of signs could begin the texts.
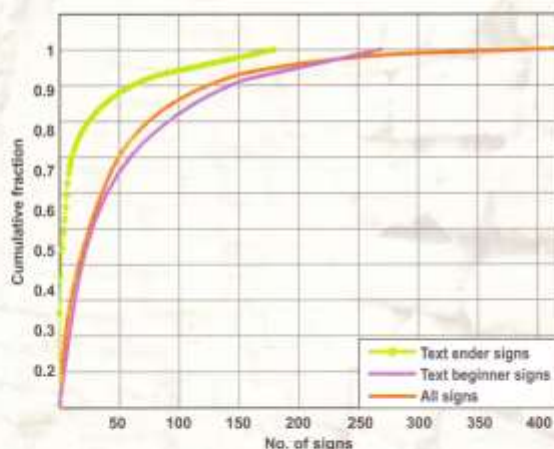
Cumulative frequency plot for all signs, text-beginners and enders [Yadav, et al., 2010].

Sequencing of signs in Indus texts: 'sign combinations of two, three and four signs appear with far higher frequency in the Indus script dataset than expected by chance.

Segmentation of Indus texts

a)   length of text in Indus script varies from 1 to 14 signs per line.
b)   text beginners and enders are well defined.
c)   88% of all texts of length five or more can be segmented into segments of length not exceeding four.
d)   multiple units of information written in a text.
e)   longer strings of writing are a collection of several smaller units of information and not a long unit of complex information.

Flexibility in sign usage across different sign systems

They consider that the results of their findings increase the evidence in favour of the linguistic hypothesis of the Indus script. It should however be noted that it does not prove it to be linguistic.

Design of Indus signs

Mahadevan records 417 distinct signs in the sign list of Indus script [Mahadevan, 1977] in 3573 lines of 2906 texts.

Several signs seem to have been designed by adding modifiers to signs or merger of several individual signs.

Types of signs:

1.   basic signs,
2.   provisional basic signs
3.   and modifiers

Sign compounding (ligaturing) and sign modification seem to change the meaning or add value to basic signs rather than save writing space.

"There is hardly any doubt that the Signs and their appearance on various objects have some linkage, this sensitivity has:"

(1) Distribution of inscribed objects:
(2) Sensitivity of the Indus script to site and type of object:
(3) Clustering of sites and types of objects:

However there exists a very strong view against the understanding that Indus signs are representing a written language/languages. It went on to analyze in details the: impossibility of the manuscript thesis, as the Indus Symbols were not even evolving in linguistic direction after at least 600 years of use. [Collapse of Indus Script Thesis: The Myth of Litrate harappan Civilization: Steve Farmer, Richard Sproat, and Michael Witzel; Electronic Journal of Vedic Studies (EJVS) 11 – 2 (2004) pp.19-57]

This small overview brings us to the clear understanding that the Indus texts have an underlying logic and syntax indicative of writing; the ordering of signs is also more rigid than random writing. There is significant asymmetry in the usage of text beginners and text enders [Kenoyer & Meadow 2010]. It leaves hardly any doubt that the manner of writing remained highly standardized over the entire civilization.

The studies on the Indus script can be fruitful when the whole context is re-questioned; more detailed analysis is bound to provide certain information that could lead to better understanding of the Indus culture.

"The intelligent writers carefully differentiated between different signs, which were clearly showing their distinct shape, and all such signs were brought together to form 'a corpus' of the Indus signs. This logically a dangerous exercise became the base of the studies of the Indus signs."

Besides the minute observations on the seals and its production have given meaningful insight in the evolution of seal forms over time. "This and similar other investigative observations must become the part of study to ascertain the possible development of signs in their form, combinations and sequencing.

The new excavations produce stratigraphically controlled and well-dated inscribed pieces, it becomes both possible and increasingly important that the script and accompanying iconography be subjected to rigorous analysis at the site, regional, and cross-regional levels to investigate trends of change through time and variability across space.

The correlation of script variations and the types of the objects bearing the signs are to be seen in context, as the sites relevant were constantly undergoing changes, transforming, expanding or and declining, Kenoyer feel that script shows the transformation as well [Kenoyer & Meadow 2010]

It was a diligent exercise that improved the description of Mahadevan when Parpola lays down clear guidelines for the recognition of basic signs, graphic variants and composite signs. The numerals are identified as a set of short stroke signs comprising up to nine strokes arranged in one or two tiers. Group of small inverted semi-circles which occur along with the 'stroke' numerals are very likely to be tens.

Experts may differ whether a given sign is basic or provisional or modifier. However Parpola has now provided the most complete documentation enabling other scholars to draw their own conclusion. [Harappan.com]

Further studies are imperative, thus there shall be no dearth of attention to the Indus script, there is always a reason to believe that further discoveries may produce some fateful material that may help assist the efforts of the scholars. However there is need that the script itself shall be given some more attention to look in to computational possibilities to help in arriving at some understanding of the context.

The Indus signs have been largely used as drawn images in computational analysis and studies. The measure has been taken to make this script as normal computer friendly font, useable as the font for future computational exercises. Thus the Indus signs are transformed in Scalable Vector Graphics (SVG) based font for installing in computers.

The Indus signs could be used in studies, in writing, as it is now available to be incorporated in writing as easily as Arabic letters or words could be typed in the text format.

Likewise these signs are now computer program, a font that could be used in computing exercises, statistical counts, as the computers easily recognize each sign.

The Indus script can be downloaded for use and computations etc. web version of this font has been developed and placed at the website of the mohenjodaroonline.net, for downloading by bloggers, researchers and web developers, free of cost.

This web version font is comprised on cascading style sheets (CSS), all versions of universal accessing fonts of the Indus Script font including tutorial of web embedding. This version of the font could be helpful to the researchers and developers universally, enabling placing these signs as text rather than the drawn images.

Technically another advantage of this web version is that the Google or other such search engines could carry it in their research algorithms.
It is being reasonable to hope that placing the script this way it shall be fully exposed to "neural networking," the computing identifier of similarities, serial sequencing and of inherent characteristics.

This feature may become instrumental further in studies by intelligent users. It may further help increase the chances of computational assistance in varied sets of reorganized reasoning, behind the use and its sequencing in the antiquity.

This font has been prepared and made available for installing on computers and embedding on websites by researchers and users around the world. It has been developed under the umbrella of NFM, by Mr. Shabir Kumbhar, engineering / embedded and mapping by Mr. Amar Fayaz Buriro, under the advise of present writer, thus any lacuna remaining may be not for want of their efforts, but shall solely be fault of this writer.

NFM-Indus Script
Unicode Private Use Area (PUA) Table of Codes

HOW TO USE THE INDUS SCRIPT FONT

For using specific sign of Indus Script font in word document, please follow these instructions:

1. Install NFM Indus Script font.

2. Find the value code of the required sign from table; vertical value & horizontal number: (For example the sign 🛉 has value E07 horizontal and 5 number in vertical. So the value code will be E075.)

3. Just type this code and then press Alt+X, without space.

4. A rectangle symbol will appear. Just select that rectangle symbol.

5. In formatting toolbar, select "NFM Indus Script" font. Required sign of Indus Script will appear at document in text format.

NFM-Indus Script Signs with the table of codes:

Value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
E00																
E01																
E02																
E03																
E04																
E05																
E06																
E07																
E08																
E09																
E0A																
E0B																
E0C																
E0D																
E0E																
E0F																
E10																
E11																
E12																
E13																
E14																
E15																
E16																
E17																
E18																
E19																
E1A																

Value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
E1B																
E1C																
E1D																
E1E																
E1F																
E20																
E21																
E22																
E23																
E24																
E25																
E26																
E27																
E28																
E29																
E2A																
E2B																
E2C																
E2D																
E2E																
E2F																
E30																
E31																
E32																
E33																
E34																
E35																

Value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
E36																
E37																
E38																
E39																
E3A																
E3B																
E3C																
E3D																
E3E																
E3F																
E40																
E41																
E42																
E43																
E44																
E45																
E46																
E47																
E48																
E49																
E4A																
E4B																
E4C																
E4D																
E4E																
E4F																
E50																

Value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
ES1																
ES2																
ES3																
ES4																
ES5																
ES6																
ES7																
ES8																
ES9																
ESA																
ESB																
ESC																
ESD																
ESE																
ESF																
E60																
E61																
E62																
E63																
E64																
E65																
E66																
E67																
E68																
E69																
E6A																
E6B																
E6C																

36

Value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
E6D																
E6E																
E6F																
E70																
E71																
E72																

STATISTICS

(The table showing statistical data of varied Indus Signs in NFM Indus Script font as identified by Asko Parpola in his compiled corpus)

NFM-Indus Script Unicode PUA Table of codes:

Value	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	Total
E00	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E10	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E20	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E30	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E40	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E50	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E60	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	256
E70	16	16	15	0	0	0	0	0	0	0	0	0	0	0	0	0	047
Total	128	128	127	112	112	112	112	112	112	112	112	112	112	112	112	112	1839