
 TIIPD

Indus Valley Civilization & its Writing

Nisha Yadav
Tata Institute of Fundamental Research, Mumbai

 TIIPD

“The practice of writing and the development of a coherent system of signs, a script, is something which is seen only in complex societies

.... Writing is a feature of civilizations.”

*-- Colin Renfrew
(Archaeology and Language, 1987)*

Indus Valley Civilization & Its Writing 2

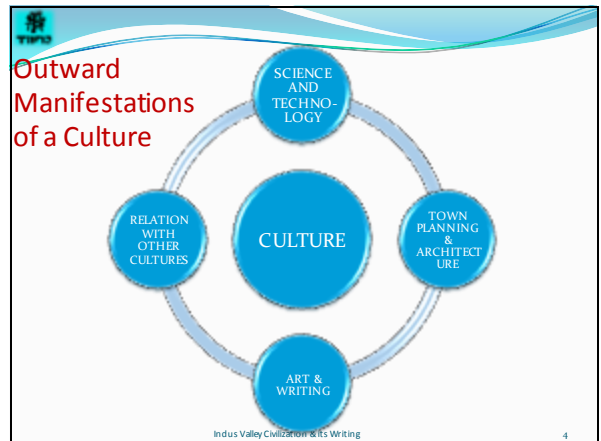
 TIIPD


Culture as a Complex system

A culture can be envisaged as a complex system for a variety of reasons:

- It is large and self-organising.
- It is multi-parameterised.
- Each parameter has certain self driven and group behaviour and the system evolution is a mix of these parameters.

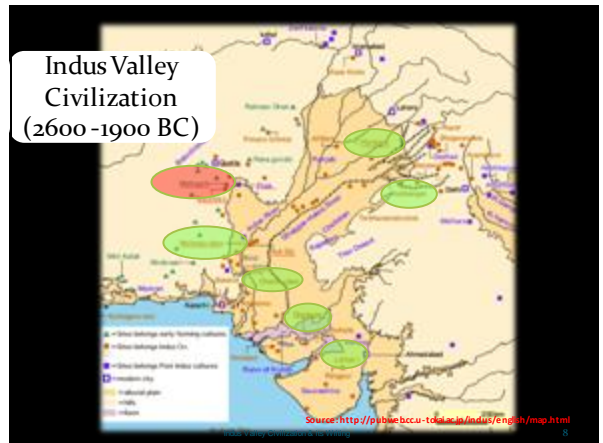
Indus Valley Civilization & Its Writing 3



 TIIPD

Study of a culture becomes more complex when the SYSTEM IS NON-DOCUMENTED and STUDIED POST-DISINTEGRATION.



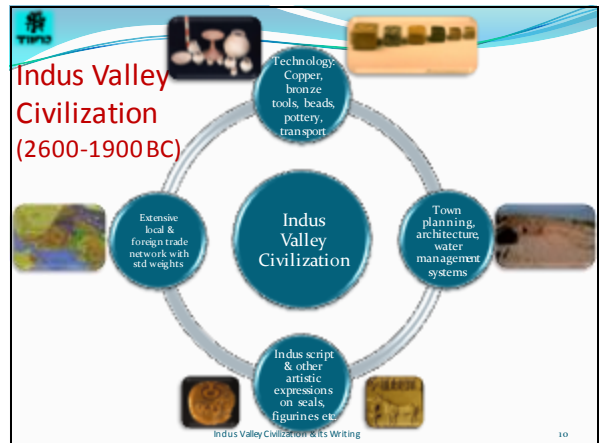


Timeline of Indus Valley Civilization

The Indus valley civilization had the following timeline:

- Roots of civilization are found in Mehrgarh around 7,000 BC.
- It went through a major expansion from 4,500 BC to about 3,200 BC.
- It evolved into an urban civilization around 2,600 BC.
- It decayed and disappeared between 1,900 BC and 1,300 BC.

Indus Valley Civilization & Its Writing 9



Present talk

In the present talk, we will discuss the following aspects of the Indus valley civilisation (IVC)

- Structure and context of IVC
- Complexity of city architecture
- Complexity of its artistic expressions
- Complexity of its script

Indus Valley Civilization & Its Writing 11



Defining Characteristics

These are:

- Seals, sealings with human or animal motifs and script
- Black – on – red pottery with design
- Long parallel – sided chert blades
- Beads of carnelian, steatite and faience
- Weights of agate
- Objects of copper and bronze
- Terracotta toys and goddess figurines
- Characteristic burials styles
- Citadels, platforms for houses
- Cotton, barley and wheat
- Sophisticated water and waste management system
- Standard brick size

Joshi (2008)

Indus Valley Civilization & Its Writing 19

General Comments

- Any civilisation as complex as this, may be assumed to have the following:
 - Commonly agreed principles of record keeping including calendar, writing, numeric systems etc.
 - Standardised trade practices including weights, currencies or barter agreements for packaging and transport.
 - Inter-affected art and cultural growth.
 - Similarity in civic practices, social behaviour and norms.
 - Agreed protocols on different habitations between urban and rural sites.
- In principle, they should also have common religious practices but we can only make conjectures about this.
- Some centralised authority that can impose its will and discipline by force – something that is conspicuous by its **ABSENCE** in IVC.

Indus Valley Civilization & Its Writing 20

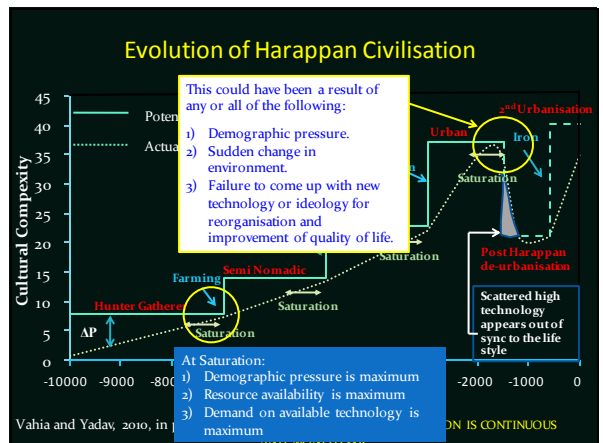
Parameterisation of IVC

BASED ON THE WORK OF MURDOCK AND PROVOST (1973)

Scale 1: Writing and Records; Scale 4: Urbanization; Scale 7: Money;
 Scale 2: Fixity of Residence; Scale 5: Tech. Specialisation; Scale 8: Density of Population;
 Scale 3: Agriculture; Scale 6: Land Transport; Scale 9: Level of Integration;
 Scale 10: Social Stratification

Stages	Scales (1 to 4, with 4 being the best)										Total
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	
Hunter-Gatherer	0	1	1	2	0	0	0	2	1	1	8
Semi-nomadic	0	2	2	2	1	1	1	2	1	2	14
Settled	1	3	3	3	2	2	1	3	2	3	23
Urban	3	4	4	4	4	3	3	4	4	4	37
Post Urban	1	3	3	2	2	2	2	2	2	2	21
2 nd . Urbanisation	4	4	4	4	4	3	4	4	4	4	39

Indus Valley Civilization & Its Writing 21



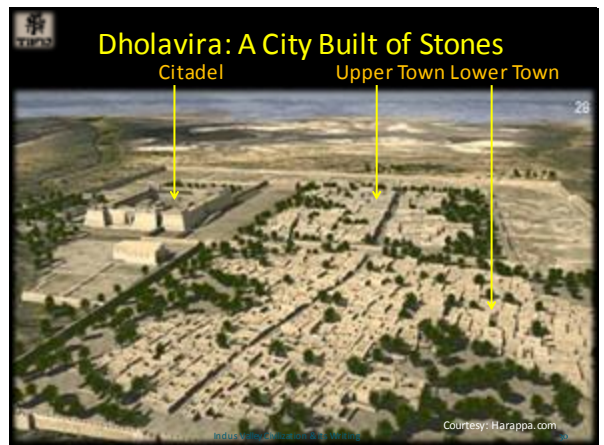
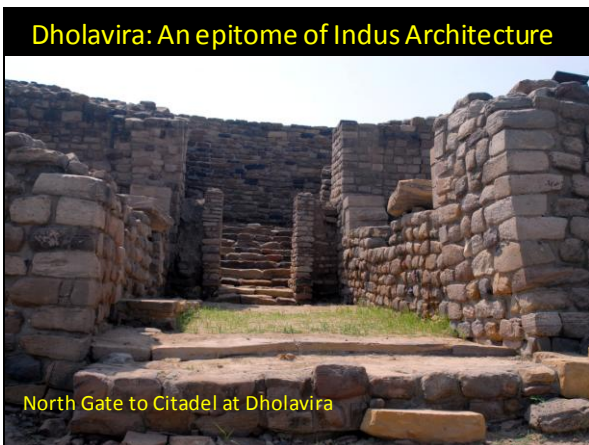
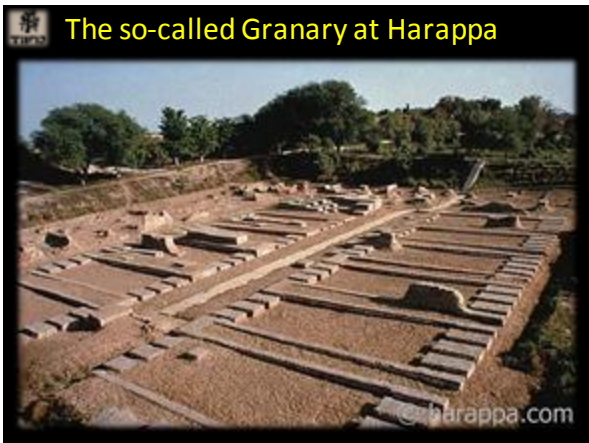
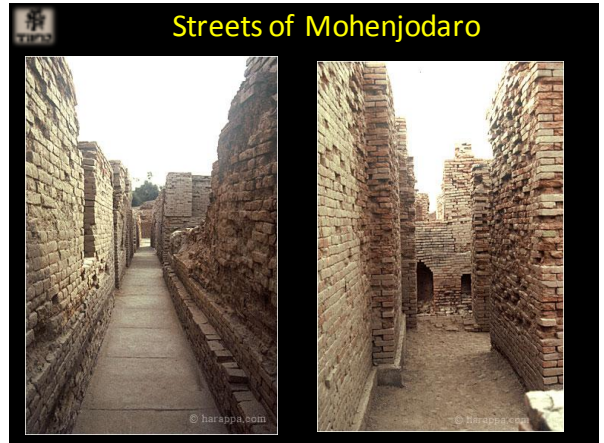
Some Conclusions

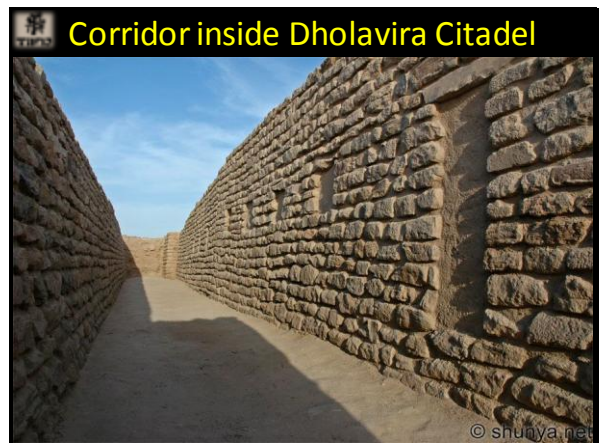
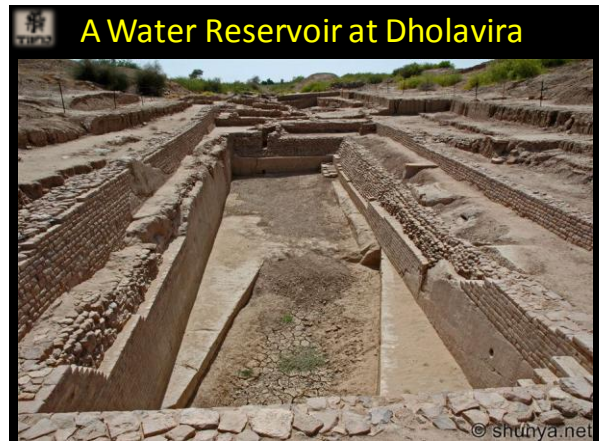
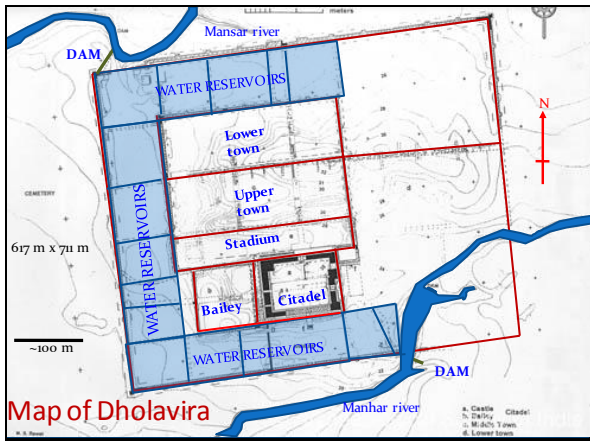
- IVC was a complex, multifaceted civilization.
- The coexistence of urban and rural lifestyles in a symbiotic manner was an important feature of the civilisation.
- Absence of grandiose structures and large standing army suggest that the civilisation was more like a Greek Poleis (but 2000 years before them) and not like the Egyptian or West Asian civilisations.
- Their social organisation and internal dynamics including stratification and interrelation between various groups was unique.

Indus Valley Civilization & Its Writing 23

Complexity of Architecture:

They had no temples or monumental buildings though the civilization had the imagination and capability.





Complexity of Architecture

- Indoor water closets and bathing facilities.
- Standardised brick usage with aesthetically designed structures.
- Several hundred meter long straight and orthogonal streets with all entries to houses that do not open in the main street.
- Long, gravity assisted water and drainage systems.
- Deep brick laid wells.

Indus Valley Civilization & its Writing 35

Complexity of its Artistic Expressions:

Obsession(?) for Miniatures!

Scale of a Typical Seal

In most cases, seals are between 2.5 to 5 square cm in size.

Photographs by Jatin Acharya

Indus Valley Civilization & its Writing 37

Components of Indus Seals

1. Script (42.30%)
2. Animal Motif (22.26%)
3. Manger (14.23%)
4. Crude (10.51%)
5. Geometry (4.55%)
6. Abstract (1.70%)
7. Scene (1.30%)
8. Human figures (1.07%)
9. Plant motif (1.01%)
10. Mythical figure (0.46%)
11. Composite animal (0.45%)
12. Multi-headed animal (0.18%)

5 cm

5 cm

Same object can have more than one of these basic components.

Analysis based on CSI Volumes 1 & 2, Yadav & Vahia (Submitted)

Indus Valley Civilization & its Writing 38

Some Interesting Components

4.0 cm

3.0 cm

3.2 cm

Courtesy: Prof. John Huntington

Indus Valley Civilization & its Writing

CSI, www.harappa.com

Some more patterns...

3.5 cm

3.4 cm

2.7 cm

2.0 cm

3.3 cm X 1.5 cm

Same object opposite sides

Relatively to scale

Courtesy: CSI, Harappa.com

Emphasis on Symmetry!

2 fold symmetry

4 fold symmetry

5 fold symmetry

7 fold symmetry

Mirror symmetry

2.4 cm

1.9 cm

2.0 cm

3.2 cm

3.5 cm

Indus Valley Civilization & its Writing

41

Geometric Seals: Pattern fitting

Basic Pattern

Basic Pattern

Basic Pattern

2 cm

2 cm

3.2 cm

Size : 4.2cm X 4.8 cm

Courtesy: CSI, Harappa.com


42



“At their very best, it would be no exaggeration to describe them as little masterpieces of controlled realism, with a monumental strength in one sense out of all proportion to their size and in another entirely related to it.”


-- Wheeler, 1968

Indus Valley Civilization & its Writing 43




Complexity of its Script:

It has defied decipherment!




What did they write?



Role of Writing in Ancient Cultures


- Writing allowed literate elites to
 - store records
 - keep a track of time
 - facilitate communication over short and long distances
 - express superiority and power
- The ambiguity associated with the context and usage of Indus signs has made the problem of Indus writing system even more challenging!

Indus Valley Civilization & its Writing 46



What adds to the challenge?

- Very short and brief texts–
 - average number of signs is 5.
 - longest single line text: 14 signs
 - longest text: 26 signs running in 3 distinct lines.
- Language(s) underneath (if any) is unknown.
- Lack of bilingual texts.



Longest single line text

Indus Valley Civilization & its Writing 47

Source: Wikipedia

Rosetta Stone



Ancient Egyptian
(Hieroglyph)

Demotic

Greek

No Rosetta stone for Indus script yet!

48

Indus Writing on Various Objects

Seals (62.16%) & Sealing (19.31%)

Miniature Tablets (9.43%)

Pottery Graffiti (2.10%)

Copper Tablet (5.37%)

Ivory sticks (0.78%)

Indus Valley Civilization & its Writing www.harappa.com, Parpola (1992)

Dholavira Sign Board

Source: pubweb.ccu-tokai.ac.jp/indus/english/2_4_03.html
Indus Valley Civilization & its Writing

The sign list of Indus script consists of about 400 to 700 signs which look like human, fish etc.

Indus Valley Civilization & its Writing

Indus Script Datasets

Database	Mahadevan (1977)	Parpola (1979)	Wells (2006)
No. of texts	3573	3700	3835
No. of signs	417	398	676
Total sign occurrences	13,772	~ 14,800	17,427

We use Mahadevan's concordance for analysis.

Indus Valley Civilization & its Writing

Number of Signs and Script Type

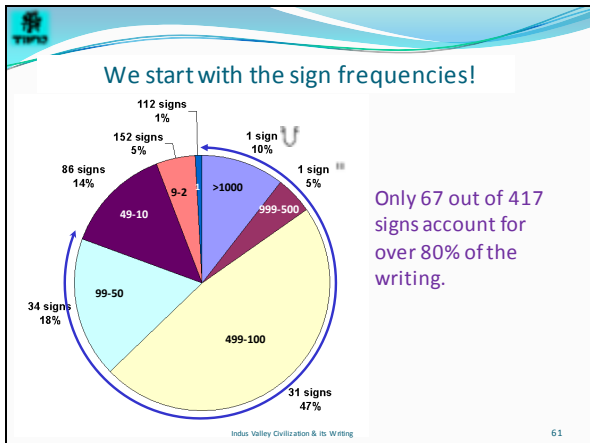
S. No	Type of Scripts	Type of Signs	No. of Signs	Examples
1.	Logographic	Word-signs	Thousands	Chinese
2.	Logo-syllabic	Word-signs & Phonetic syllables	900-400	Sumerian, Egyptian
3.	Syllabic	(a) Closed & Open syllables	200-100	Elamite, Cuneiform
		(b) Open syllables	100-40	Linear B, Old Persian
4.	Alphabetic	Single-sound signs	Below 40	Semitic, Greek, Latin

Indus Valley Civilization & its Writing

Indus Script Signs (1 to 110)

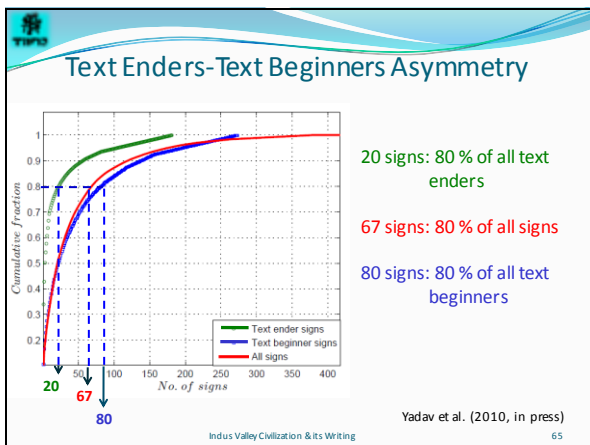
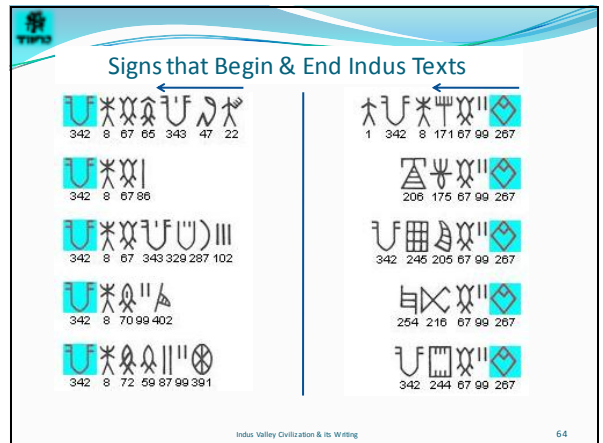
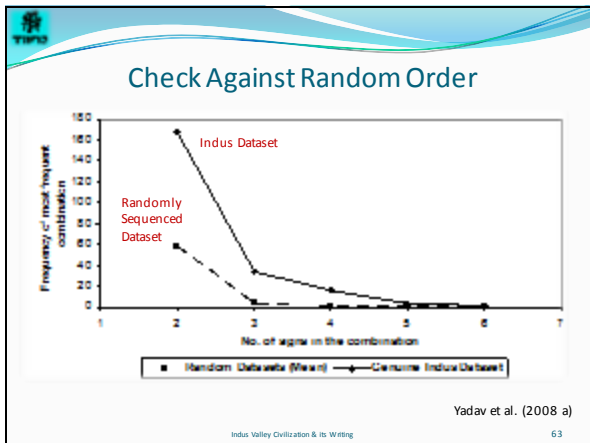
From Mahadevan (1977)

Indus Valley Civilization & its Writing



- The first question we wanted to answer was: "Are Indus texts just a collection of randomly ordered signs or is there any sequencing?"
- So we randomized all the written material and checked how often we get similar sequences by chance and in real Indus data.

Indus Valley Civilization & Its Writing 62



Where do these frequent sign sequences occur in Indus texts?

Positional Analysis of Frequent Two-sign Combinations

Two-sign Combination	Frequency	Left (%)	Middle (%)	Right (%)	
99	267	168	1.79	11.90	85.71
89	336	75	0.00	89.33	10.67
176	342	59	96.61	3.39	0.00
342	8	58	72.41	25.86	0.00
99	391	56	0.00	8.93	91.07
342	347	56	89.29	10.71	0.00
1	342	48	89.58	10.42	0.00
123	293	40	0.00	0.00	100.00
59	87	39	0.00	79.49	20.51
342	48	38	52.63	28.95	15.79
59	171	36	0.00	80.56	19.44
162	249	34	0.00	85.29	14.71
211	89	34	91.18	8.82	0.00
245	33	33	60.61	21.21	18.18
211	59	31	90.32	9.68	0.00
67	65	27	0.00	74.07	25.93
130	51	27	7.41	70.37	22.22
67	99	26	0.00	100.00	0.00
342	162	25	84.00	12.00	0.00
343	123	25	0.00	100.00	0.00

Indus Valley Civilization & Its Writing | Yadav et al. (2008 a) 67

Positional Distribution of Sign Sequences

85 % of times starts a text.

89 % of times comes in the middle of texts.

96 % of times ends a text.

Indus Valley Civilization & Its Writing | Yadav et al. (2008 a) 68

Positional Analysis of Frequent Triplets

Three-sign Combination	Frequency	Left (%)	Middle (%)	Right (%)		
211	89	336	34	88.24	5.88	2.94
343	123	293	25	0.00	0.00	100.00
342	162	249	24	83.33	8.33	4.17
342	169	249	20	70.00	20.00	5.00
342	8	171	19	73.68	5.26	15.79
149	130	51	19	0.00	78.95	21.05
59	87	99	16	0.00	100.00	0.00
342	87	403	16	81.25	6.25	6.25
342	149	130	16	75.00	25.00	0.00
67	99	267	14	0.00	7.14	92.86
87	99	267	14	0.00	21.43	78.57
89	336	72	14	0.00	85.71	14.29
65	99	267	12	0.00	8.33	91.67
342	244	67	12	66.67	8.33	16.67
15	389	178	11	72.73	0.00	18.18
59	171	53	10	0.00	60.00	40.00
245	245	25	10	90.00	0.00	0.00

Indus Valley Civilization & Its Writing | Yadav et al. (2008 a)

Can we say that longer texts have multiple units strung together?

Segmentation of Indus Texts

By comparing longer and shorter texts on different objects!

Mahadevan (1986), Yadav et al. (2008b)

Indus Valley Civilization & Its Writing | 71

Segmentation of Indus Texts

By using frequent sign sequences!

Mahadevan (1986), Yadav et al. (2008b)

Indus Valley Civilization & Its Writing | 72

Segmentation of Indus Texts

By comparing adjacent pair frequencies!

Mahadevan (1986), Yadav et al. (2008b)

Indus Valley Civilization & Its Writing 73

Indus Dataset Before and After Segmentation

Dataset before Segmentation

Dataset after Segmentation

Text lengths 1 to 14

No segment > 9 signs

Yadav et al. (2008b)

Indus Valley Civilization & Its Writing 74

Examples of Segmentation

Object No.	Segments of Text						
1232:	P148	P96					
1279	4441						
4254:	P53	T140	P116	P49	309		
2371	2015	1226					
2637:	P41	PM14	67	PM9	309	344	PB1
	6001		1093				495
2461:	T34	326	87	P131	178		
1437	2673	4560	2682				

Indus Valley Civilization & Its Writing Yadav et al. (2008b)

Network Analysis

Modular partitioning of the network

Modularity optimization \Rightarrow 8 communities (shown as differently colored nodes)

Sinha, Pan, Yadav et al. (2009, ACL-IJCNLP)

Indus Valley Civilization & Its Writing 75

Results from Network Analysis

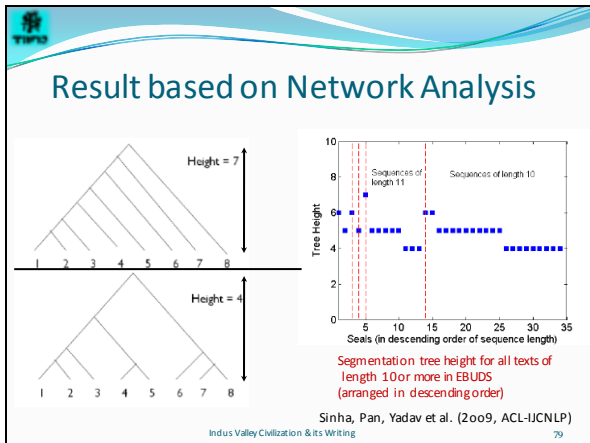
distribution for 10^4 realizations of corpus randomization

EBUDS

Probability of occurrence of the sign pair 267-59

Sinha, Pan, Yadav et al. (2009, ACL-IJCNLP)

Indus Valley Civilization & Its Writing 76



Analysis using N-gram & Markov Models

- ### Applying Machine Learning and Data Mining
- Research in machine learning and data mining has led to new techniques for:
 - Learning statistical models of sequences
 - Grammar discovery
 - Pattern recognition
 - Pattern completion
 - We apply these new techniques to the Indus script problem.
- Indus Valley Civilization & its Writing 81

- ### N-gram & Markov Models
- Are probabilistic models which provide a very useful method of modelling different types of sequences.
 - These models are not sensitive to the semantic content of the sequences but, reveal the syntax, if any, that the sequences follow.
 - The order of the Markov model decides the length of correlation.
- Yadav et al. (2010, in press), Rao, Yadav et al. (2009, PNAS)
- Indus Valley Civilization & its Writing 82

Markov Model for 3 States A, B and

A	0.2	0.8	0
B	0.3	0.7	0
#	0	0	1
	A	B	#

Some example Sequences with this Markov model: BAAB, ABAB, B, etc. (the terminal sign # is not shown).

Sequences not seen with above Markov model: All texts with repetition of BB or all texts which end in A.

Rao, Yadav et al. (2009, PNAS)

Indus Valley Civilization & its Writing 83

How can we model English alphabet in bigram?

$S = \{A, B, C, \dots, M, N, \dots, Z\}$ → 26 States

sign	A	B	...	Z
A	$p_{1,1}$	$p_{1,2}$...	$p_{1,26}$
B	$p_{2,1}$	$p_{2,2}$...	$p_{2,26}$
...
Z	$p_{26,1}$	$p_{26,2}$...	$p_{26,26}$

→ Transition matrix (26 × 26)

$\Pi = \{p(X_0 = A), p(X_0 = B), \dots, p(X_0 = Z)\}$ → Ini. state prob. vector

Indus Valley Civilization & its Writing 84


Examples from English

- At the level of individual word, with each *letter* as an independent token, we find that
 - the letter 't' can be succeeded by lot of letters such as 'a', 'e', 'o', but not 'x' or 'z',
 - the letter 'q' usually succeeded by 'u' and so on.
- Similarly, if we the tokens are *words*, then we find that the word 'the' can be succeeded by a large number of words, but not verbs such as 'eat'.

Indus Valley Civilization & its Writing 85

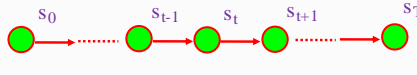
Indus Texts are mapped to sequences

Indus Text



M-355 A

Sequence of states



Yadav et al. (2010, in press); Rao, Yadav et al. (2009, PNAS)
Indus Valley Civilization & its Writing 86

How do we model the Indus Script?

$S = \{ \text{𑀓}, | \text{𑀓}, \text{𑀓}, \dots, \text{𑀓}, \dots, \text{𑀓}, \dots, \text{𑀓} \}$ → 417 States

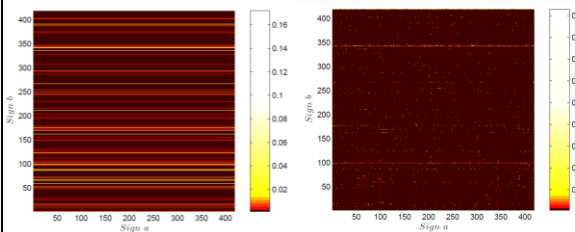
sign	𑀓	𑀓	...	𑀓
𑀓	$P_{1,1}$	$P_{1,2}$...	$P_{1,417}$
𑀓	$P_{2,1}$	$P_{2,2}$...	$P_{2,417}$
...
𑀓	$P_{417,1}$	$P_{417,2}$...	$P_{417,417}$

→ Transition matrix (417 × 417)

$\Pi = \{ p(X_0 = \text{𑀓}), p(X_0 = | \text{𑀓}), \dots, p(X_0 = \text{𑀓}) \}$ → Initial state prob. vector

Indus Valley Civilization & its Writing 87

Comparison of Bigram Matrices



No correlation Indus Script

Yadav et al. (2010, in press)
Indus Valley Civilization & its Writing 88

Markov Model of Indus Texts

We use this Markov model of Indus texts for

- filling-in damaged or illegible Indus texts
- generating Indus like texts
- finding the likelihood of a string to the learned model - how closely the statistical properties of a string matches the texts used for learning the model?



Yadav et al. (2010, in press); Rao, Yadav et al. (2009, PNAS)
Indus Valley Civilization & its Writing 89

a) Filling in damaged & illegible texts



Indus Valley Civilization & its Writing 90

Comparing West Asian vs. Indus Area Inscriptions: Difference in structure?

Synthetic example:

Indus text		Likelihood	2.8×10^{-5}
Altered text			~ 0

Actual example:

Indus seal			1.4×10^{-6}
West Asian seal			5.6×10^{-11}

West Asian seal text is approximately 100,000 times as unlikely to be generated by the learned model as other Indus area texts.

Rao, Yadav et al. (2009, PNAS)

Indus Valley Civilization & Its Writing 97

All this shows that

- Indus writing was clearly ordered and seems to be dictated by certain set of rules.
- These results seem to suggest an underlying grammar.
- Does the script represent language?

Indus Valley Civilization & Its Writing 98

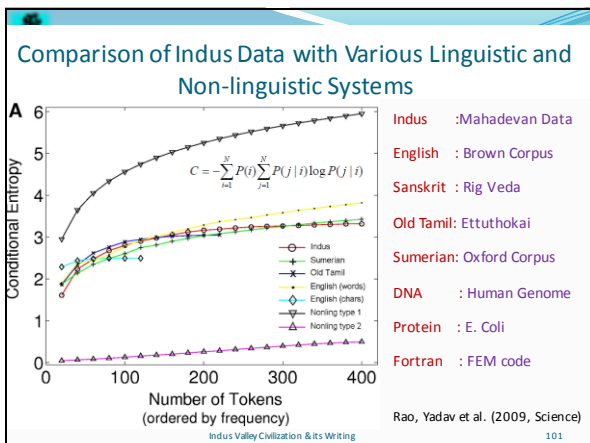
Comparison with other Linguistic & Non-linguistic Systems

Comparison with Other Sign Systems

- We have compared the conditional entropy of Indus sign system with other linguistic and non-linguistic sign systems.
- The study *indicates* that the flexibility in choosing a sign given a preceding sign in Indus script is relatively close to linguistic systems than non-linguistic sign systems.

Rao, Yadav et al. (2009, Science)

Indus Valley Civilization & Its Writing 100



Summary

- Harappan civilization was spread over 1.5 million km².
- It had a complex network of cities and villages that constituted the civilization.
- Its growth and decay can be evaluated using its standard of living.
- Its architecture and technology were also highly evolved.
- Its writing is an excellent example of its creativity.
- Study of Indus writing suggests that the script has a rich syntax with an underlying logic in its structure. Its meaning is not known to us.
- We can't read Indus script but (we think) we can WRITE it with some degree of confidence.

Indus Valley Civilization & Its Writing 102



Collaborators

- Mayank Vahia, TIFR, Mumbai
- Iravatham Mahadevan, IRC, Chennai
- Rajesh Rao, University of Washington, Seattle
- Hrishikesh Joglekar, Oracle, Hyderabad
- Ronojoy Adhikari, IMSc, Chennai
- Sitabhra Sinha, IMSc, Chennai
- Raj K Pan, IMSc, Chennai

Acknowledgement

- <http://www.harappa.com>

Our published work is available at:

- <http://www.harappa.com>
- <http://www.indusresearch.wikidot/script>

Indus Valley Civilization & Its Writing 103





Our Papers

1. N. Yadav, et al., *A statistical approach for pattern search in the Indus writing*, International Journal of Dravidian Linguistics, **37**, 39, 2008.
2. N. Yadav et al. *Segmentation of Indus texts*, International Journal of Dravidian Linguistics, **37**, 53, 2008.
3. M. N. Vahia and N. Yadav, *Harappan Weights*, Puratattva, Vol. 32, 2008.
4. R. P. N. Rao, N. Yadav et al., *Entropic evidence for linguistic structure in the Indus script*, Science, **324**, 1165, 2009.
5. R. P. N. Rao, N. Yadav et al., *A Markov model for the Indus script*, Publications of the National Academy of Sciences, 2009.
6. N. Yadav et al., *Statistical analysis of the Indus script using n-grams*, submitted for publication, available at arxiv.org, 2009
7. S. Sinha, Pan, N. Yadav et al., *Network analysis reveals structure indicative of syntax in the corpus of undeciphered Indus civilization inscriptions*, ACL-IJCNLP 2009, pp. 5-13, 2009
8. N. Yadav and M. N. Vahia, *Classification of patterns on Indus objects*, Submitted to IJHS, June 2009
9. M. N. Vahia and N. Yadav, *Harappan Geometry and Symmetry: A study of geometrical patterns on Indus Objects*, Submitted to IJHS, June 2009

Our published work is available at:

- <http://www.harappa.com> <http://www.indusresearch.wikidot/script>

Indus Valley Civilization & Its Writing 104

Thank you!

Indus Valley Civilization & Its Writing 105