FIFTY FIRST ANNUAL MEETING OF THE INDIAN ACADEMY OF SCIENCES, HELD AT MADURAI KAMARAJ UNIVERSITY CAMPUS, MADURAI. SUMMARIES OF LECTURES DELIVERED

A Computer Study of the Indus Script by I. Mahadevan, *International Association of Tamil Research, Madras. (Residence: Vyjayanthi, 112, Chamiers Road, Nandanam, Madras 600 035, India).*

This paper is a brief report of the ongoing programme of a computer study of the Indus Script carried out by the present author and Mythili Ranga Rao at the Tata Institute of Fundamental Research, Bombay. A Database for the Indus Script has been compiled on the CYBER 170/730 Computer System. Preliminary results of the statistical analysis indicate that none of the claims of decipherment of the Indus Script made so far is wholly successful.

1. *Database for the Indus Script*

A Database for the Indus Script has been compiled by the authors on the CYBER 170/730 Computer System at the TIFR, Bombay. The database consists presently of 3573 lines of text found written on 2906 artefacts from 19 Harappan and 5 West Asian sites. The artefacts (designated as 'inscribed objects') have been classified broadly into eight types based on the materials of manufacture and modes of writing. A summary of the distribution of the inscribed objects according to sites and types is given in table 1.

The database has been divided into two parts namely, background data and the Texts. The background data provide information on (i) site, (ii) locus of occurrence within the site, (iii) Stratigraphic level, (iv) object type, (v) associated field symbols (pictorial motifs), (vi) direction of writing of each line of text and (vii) the number of signs in each line. A reference number, which is uniquely defined, is assigned to each line of text. The Texts reproduce the Indus Signs occurring in the inscriptions in both numerical and literal (pictorial) forms to facilitate statistical study and reporting respectively. A brief description of the elements of the database is given in table 2. The database also includes a library of signs in the Indus Script for graphic reproduction on a CALCOMP Drum Plotter (examples in figure 1).

Table 1 *Distribution of inscribed objects according to types and sites*

Types of inscribed Objects	No. of Occurrences at Sites							
	MD	HP	CD	LL	KB	OS	WA	CORP. Total
Seals	1232	350	58	89	56	13	16	1814
Sealings	119	288	3	75	21	4	1	511
Miniature Tablets	–	272	–	–	–	–	–	272
Pottery Graffiti	13	64	4	1	20	17	–	119
Copper Tablets	135	–	–	–	–	–	–	135
Bronze implements	5	3	1	–	2	–	–	11
Ivory/bone rods	28	1	–	–	–	–	–	29
Misc. Obj.	8	7	–	–	–	–	–	15
Total	1540	985	66	165	99	34	17	2906

[Notes: MD: Mohenjodaro, HP: Harappa, CD: Chanhudaro, LL: Lothal, KB: Kalibangan, OS: Other sites, WA: West Asian Sites, CORP.: Corpus of Texts. Misc. Obj: Miscellaneous Inscribed objects. See statements 1 and 2 for analysis.]

Table 2 *Data description of the fields in the data base*

Attributes	Description
Reference Number	Each line of text has a unique Ref. No. in 6 digits comprising the Site Number (col. 1), the Object Number (cols. 2–4), the Side Number *i.e.* the number of inscribed faces of the object (col. 5) and the Line Number, *i.e.* the number of lines of text on each inscribed side of the object (col. 6).
Locus	Area, Section or sub-section of the site as determined by the excavator.
Level	The Level in ft. at which the object was found above (+) or below (−) the datum (in Mackay's excavations at Mohenjodaro and Chanhudaro), or below (−) the surface in Marshall's excavation of Mohenjodaro and Vats' excavation of Harappa. (The data on levels on other sites are not available.) The levels are rounded off to the nearest foot.
Type	The typology of the inscribed objects. (See Table 1 for list of types)
Field Symbol	The pictorial motif in the field on each side of the inscribed object. (See I. Mahadevan 1977, pp. 793–813 for the list of field symbols and illustrations)
Direction of Writing	Mostly from right, occationally from the left and rarely from top to bottom. (The direction of writing was determined by the criteria discussed in I. Mahadevan 1977, pp. 10–14.)
No. of positions in a Line of text	This number records the total number of signs and text breaks (or illegible portions) in a line of text for computational processes.
No. of Signs in a Line of Text	This number indicates the total of extant and legible signs in a line of text.
Line of Text	Each line of text is coded as a series of 3-digit numbers each uniquely defining a sign. (For the Sign List of the Indus Script, see I. Mahadevan 1977, pp. 32–35). Doubtful signs are marked by asterisks. Breaks and illegible portions are also indicated by a special symbol.

Note: The Corpus of Texts published by I. Mahadevan (1977) is based on this Input Data; but the format in the book has been slightly re-arranged. Data on Locus and Level, and the number of 'positions' and signs have been omitted, the Field Symbol codes abridged.

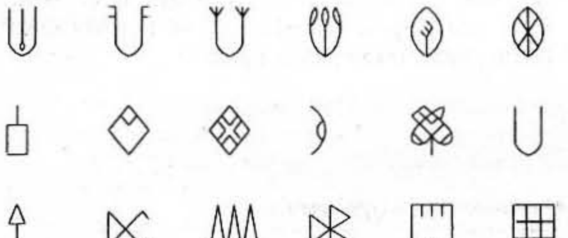

Figure 1. Specimens from the Library of Indus Signs produced on Calcomp Plotter

2. *Application of the database: Preliminary results*

A Corpus of Texts, Sign-Concordance and Statistical Tables (in an abridged form) compiled from the database have already been published[1]. (For another independent effort to reproduce the Indus Texts and compile a Sign-Concordance with the aid of computer, a reference may be made to the publications of a Finnish Group[2]). The algorithm to prepare the concordance, and methods to determine the direction of writing and for segmentation of texts into probable 'words' and 'phrases' have also been described by the authors in earlier papers[3,4].

Currently work is being carried out in the following areas:— (a) Study of the inscriptions in their archaeological context of occurrence[5], (b) Formal analysis of the texts to recognise grammatical features of the language[6] and (c) Evaluation of proposed claims of decipherment by trying to match the known frequency-distribution characteristics of the Indus signs with the features of the script and the language proposed by would-be decipherers[7]. None of the claims tested so far has come out successfully. But it should be emphasised that the tests are largely negative in character, serving to eliminate the unlikely but unable to pick out the uniquely correct solution!

1. Mahadevan, I., *The Indus Script: Texts, Concordance and Tables*, Archaeological Survey of India, New Delhi, 1977.
2. Koskenniemi, K. and Parpola, A., *Corpus of Texts in the Indus Script; A Concordance to the texts in the Indus Script*, Research Reports 1 and 3, University of Helsinki, 1979, 1982.
3. Rao, Mythili, R. and Mahadevan, I., *Computerised Concordance to the Texts in the Indus Script*, Paper read at the III Int. Conf. on Computing in Humanities, University of Waterloo, Canada, 1977.
4. Mahadevan, I., Recent Advances in the study of the Indus Script, *Puratattva*, 9 (1977–8), New Delhi, 1980, p. 34.
5. Rao, Mythili, R. and Mahadevan, I., *Archaeological Context of the Indus Texts at Mohenjodaro*, paper read at the Int. Conf. on Databases in the Humanities and Social Sciences, Grinnell College, Iowa, June, 1985.
6. Mahadevan, I., *Towards a Grammar of the Indus Texts*, paper read at the Seminar on Indus Script, Tamil University, Thanjavur, Sept. 1983.
7. Mahadevan, I., *Claims of decipherment of the Indus Script: Some objective methods to test their validity*, paper read at the South Asian Regional Conf. (SARC) Workshop on Epigraphy, Mysore, March 1985.

Computer Engineering

Fifth Generation Computers by V. Rajaraman, *Computer Centre, Indian Institute of Science, Bangalore 560 012, India.*

INTRODUCTION

The term *Computer Generation* has been widely used most often in relation to the electronic devices used to construct computers. We will, however, take a broader perspective and consider, besides electronic devices used, the storage technology used, the mean time between failures (MTBF), the evolution of software and applications through the various generations. In the first generation the devices used were vacuum tubes. The memory was made up of acoustic delay lines and later magnetic drums. Tubes were not very reliable and consequently the mean time between failures of the computer was about an hour. The machines were programmed using machine codes or assembly languages which were one to one transformation of machine code. Early applications were mostly for scientific and engineering calculations. A big revolution took place in the second generation of computers due to the advent of transistors. Reliability of computers dramatically increased. This coupled with the invention of magnetic core memories made the use of high level machine independent languages such as FORTRAN feasible. This led to a rapid growth in computer usage in both science and business. The next step was more evolutionary. Transistors were replaced by integrated circuits with a consequent reduction in cost and ten fold increase in reliability. High level languages improved and were standardised by international standards organization. From a user's point of view a major advance was time sharing of computers with consequent interactive use of computers. The interactive use led to tremendous increase in the productivity of users as they could build models in close association with the computer. Currently we are in the fourth generation which is characterised by the microcomputer revolution. Large number of microprocessors are used in personal computers, in sophisticated instruments and even in household appliances. Other major characteristics of this generation are the advent of graphics and of computer networks. Table 1 is a comparative summary of the characteristics of the four generations of computers.

VON NEUMANN ARCHITECTURE

Computers have become faster, smaller and cheaper but one fundamental feature has remained invariant throughout the four generations. This is the basic logical structure of the computer which was proposed by Von Neumann in 1946. This logical structure also known as *architecture* is characterised by the program being stored in a central main memory and interpreted sequentially by the central processing unit. Another central concept in this architecture is the identification of a variable by a storage location which gets updated by program overwriting the contents. As a consequence of this logical structure of the machine, mostly sequential algorithms have been designed. Many algorithms are characterised by extensive use of loops, which are inherently sequential and involve dynamically changing the state of the memory. The correctness of algorithms is difficult to establish as these algorithms do not satisfy many universally understood mathematical laws. Computer languages are designed to efficiently represent algorithms and consequently are sequential languages.