INITIATIVE FOR CUNEIFORM ENCODING

JOHNS HOPKINS UNIVERSITY

NOVEMBER SECOND & THIRD, 2000

The Initiative for Cuneiform Encoding (ICE) is an international group of cuneiformists, Unicode experts, software engineers, linguists, and font vendors organized for the purpose of proposing a standard computer encoding for Sumero-Akkadian cuneiform, the world's oldest attested writing system.

The first ICE conference will be held at The Johns Hopkins University, Baltimore, Maryland, USA, on November 2nd and 3rd, 2000.

The main goals of the conference are to discuss the theoretical and practical issues surrounding the encoding of cuneiform and to establish a working group to develop a formal cuneiform encoding proposal to the Unicode Consortium.

READING CUNEIFORM

Life-size, copper statue of King Naram-Sin
Old Akkadian inscription (ca. 2230 BC)
First attested claim to deification
Sumer 1976, v32, p75a

This is what needs encoding

There are two main phases to reading a cuneiform text, the descriptive and the interpretive.

In the descriptive phase, one establishes what cuneiform signs, or graphemes, are actually written, i.e. the text is transcoded.

In the interpretive phase, one strives to make sense of the transcoding. This is a complex, iterative process which includes grapheme sequencing, transliteration, phonemic normalization, and translation.

DESCRIPTIVE

URU	NI
KI	NI
SU	
NU	IŠ

INTERPRETIVE

NI NI IŠ URU KI SU NU
ì - lí - iš URU^{KI} - šu - nu
iliš ališunu
as the god of their city

Sumero-Akkadian cuneiform, attested by hundreds of thousands of documents in many genres and several languages from various cultures spanning three millennia, is a complex syllabographic and logographic script system with perhaps a couple thousand distinct graphemes (characters). It is marked by extensive multi-valency - one grapheme can have multiple phonemic and semantic realizations.

To this day cuneiform lacks a standard computer encoding. The general practice among cuneiformists of working almost exclusively in Roman alphabetic transliteration, although suitable for its intended purposes, presents difficulties for the application of computers to cuneiform research and instruction.

The simple addition of graphemically encoded cuneiform to the current practice of transliteration will enable a dramatic increase in philological and linguistic productivity.

For example, with cuneiform encoded:

• One could easily search for cuneiform plain text in a mixed script environment, something that is practically impossible in transliteration. One could, for example, search an electronic Chicago Assyrian Dictionary , or the web, for cuneiform content.

• One could do context-free text processing of cuneiform - such as automated character recognition of cuneiform tablets (cuneiform OCR) and proximity analysis of grapheme patterns.

• Font vendors would have much greater incentive to create the many large and complex font sets needed for rendering cuneiform usefully.

Deborah Anderson, Visiting Scholar, Linguistics, Univ. of California at Berkeley | Lloyd Anderson, Linguist , Font Vendor, Ecological Linguistics | Carl-Martin Bunz, M.A., Indo-European Linguist, University of Saarland, Germany| Jerrold Cooper, Prof. of Assyriology & Sumerian, Johns Hopkins | Robin Cover, SGML/XML, Oasis | T. R. Davis, Lecturer in Bibliography and Palaeography, Univ. of Birmingham, England | Patrick Durusau, Dir. Research and Development, Society of Biblical Literature, Emory Univ. | Robert Englund, Prof. of Assyriology & Sumerian, UCLA | Michael Everson, co-owner Everson Gunn Teoranta, Ireland | Karljürgen Feuerherm, PhD Candidate in Akkadian, Computer Scientist, Univ. of Toronto | Gene Gragg, Prof. of Near Eastern Languages & Linguistics, Dir. Oriental Institute, Univ. of Chicago | Edwin Hart, Senior Computing Staff, Applied Physics Laboratory, Johns Hopkins | Thomas Izbicki, Near Eastern Resource Services Librarian, Johns Hopkins | John Jenkins, System Software Engineer, Apple, Unicode Technical Director | Charles Jones, Research Associate - Bibliographer, Oriental Institute, Univ. of Chicago | Alasdair Livingstone, Reader in Assyriology, Univ. of Birmingham, England | John McGinnis, PhD Cambridge, England | Rick McGowan, System Software Engineer, Apple | Piotr Michalowski, Prof. of Assyriology, Univ. of Michigan | David Owen, Prof. of Ancient Near Eastern & Judaic Studies, Cornell Univ. | Simo Parpola, Prof. of Assyriology, Univ. of Helsinki | Gonzalo Rubio, Asst. Prof of Assyriology, Ohio State Univ. | Dean A. Snyder, Software Engineer, Semitic/Classics Philologist, Johns Hopkins | Steve Tinney, Prof. of Assyriology & Sumerian, Univ. of Pennsylvania | Lee Watkins, Jr., Dir. Research & Instructional Technologies, Johns Hopkins | Bruce Wells, PhD Candidate in Akkadian, Johns Hopkins | Kenneth Whistler, Software Engineer, Sybase, Unicode Technical Director, Managing Editor The Unicode Standard, Version 3.0. The conference was made possible by the generous support of: Sayeed Choudhury, Digital Knowledge Center, Milton S. Eisenhower Library, Johns Hopkins | James Neal, Director, Milton S. Eisenhower Library, Johns Hopkins | Gary Ostrander, Associate Dean of Research, Krieger School of Arts and Sciences, Johns Hopkins | Stephanie Reel, CIO and Vice-Provost of Information Technology, Johns Hopkins | Lee Watkins, Jr., Director, Research & Instructional Technologies, Johns Hopkins | Johanna Zacharias, Director, Communications, Krieger School of Arts and Sciences, Johns Hopkins

ICE
BECAUSE IT'S COOL
INITIATIVE FOR CUNEIFORM ENCODING

For information contact Dean A. Snyder at 410 961-8943 or dean.snyder@jhu.edu