

The Digital Hammurabi Project

Lee Watkins, Jr. and Dean A. Snyder
The Johns Hopkins University, USA
<http://www.jhu.edu/digitalhammurabi/>

[All 3D scans of cuneiform tablets, unless otherwise noted, are provided by the Visual Information Technology Group of the National Research Council of Canada, <<http://www.vit.iit.nrc.ca/VIT.html>>.]

Abstract

The Digital Hammurabi Project, originating at The Johns Hopkins University, is a multi-disciplinary effort aimed at addressing the two major technological obstacles to greater productivity in cuneiform research - the lack of a standard computer encoding for cuneiform text and the current inadequate state of the graphic representation of cuneiform tablets.

Building on an initial 3-year, \$1.6 million grant from the U.S. National Science Foundation, the goals of the project are to:

- 1) design and produce a portable, very high resolution, very fast, 3D surface scanner for cuneiform tablets;
- 2) develop new computer algorithms to manipulate and render gigabyte-size virtual tablets in real-time over the Internet;
- 3) coordinate a formal proposal to the Unicode Consortium for a standard Sumero/Akkadian cuneiform computer encoding;
- 4) collaborate in the development of international XML standards for 3D data, cuneiform bibliography, and cuneiform text markup;
- 5) establish a petabyte-scale digital library and museum of virtual 3D cuneiform tablets;
- 6) invent a completely new technology - automated 3D character recognition of cuneiform writing.

The 16-member Hopkins team is made up of faculty, staff, and students drawn from several divisions of the university, including the Department of Near Eastern Studies, the Applied Physics Laboratory, the Department of Computer Science, and the Sheridan Libraries.

Keywords: 3D cuneiform, cuneiform encoding, digital museum, virtual cuneiform tablets, 3D surface scanning, Unicode

Introduction

Cuneiform is the world's oldest known writing system. The earliest texts appear in Mesopotamia around 3200 BC and the last native cuneiform texts were written around 75 AD. The ancient scribes pressed the ends of their styluses into damp clay in order to write the approximately 800 different logographic, syllabic, or alphabetic signs. The clay tablets were sun-dried or oven-baked and were thus preserved in the sands of the Near East for millennia.

Since the decipherment of Babylonian cuneiform some 150 years ago museums have accumulated perhaps 300,000 tablets written in most of the major languages of the Ancient Near East - Sumerian, Akkadian (Babylonian and Assyrian), Eblaite, Hittite, Persian, Hurrian, Elamite, and Ugaritic. These texts include genres as variegated as mythology and mathematics, law codes and beer recipes. In most cases these documents are the earliest exemplars of their genres, and cuneiformists have made unique and valuable contributions to the study of such modern disciplines as history, law, religion, linguistics, mathematics, and science.

In spite of continued great interest in mankind's earliest documents it has been estimated that only about 1/10 of the extant cuneiform texts have been read even once in modern times. There are various reasons for this: the complex Sumero/Akkadian script system is inherently difficult to learn; there is, as yet, no standard computer encoding for cuneiform; there are only a few hundred qualified cuneiformists in the world; the pedagogical tools are, in many cases, non-optimal; and access to the widely distributed tablets is expensive, time-consuming, and, due to the vagaries of politics, becoming increasingly difficult.

The Digital Hammurabi Project, originating at the Johns Hopkins University, is a multi-disciplinary effort aimed at addressing the two major technological obstacles to greater productivity in cuneiform research - namely, the lack of a standard computer encoding for cuneiform text and the current inadequate state of the graphic representation of cuneiform tablets.

Computer Encoding of Cuneiform

Cuneiform research is hampered by the lack of a standard computer encoding for cuneiform text, there being no ASCII equivalent for cuneiform. Instead we have multiple, competing, and proprietary encodings. This is a strong disincentive for producing the kinds of sophisticated and complicated text analysis software and fonts needed for greater productivity in cuneiform philology.

Continuing the work of the Initiative for Cuneiform Encoding, an international group of cuneiformists, Unicode experts, software engineers, linguists, and font architects (<<http://www.jhu.edu/ice>>), we are facilitating an encoding proposal to the Unicode Consortium for Sumero/Akkadian cuneiform, to include characters from the Sumerian, Akkadian, Eblaite, Hittite, Elamite, and Hurrian languages, but not Old Persian. (A loosely related proposal for alphabetic Ugaritic cuneiform has already been accepted for Unicode consideration.)

Graphic Representation of Cuneiform

Cuneiform, as written by ancient scribes, is, by its very nature, a three dimensional writing system:

- * most of the media are three dimensional - rounded clay tablets;
- * the characters themselves are three dimensional - stylus tracings and wedge impressions in wet clay;
- * the writing streams are three dimensional - sentences often run over the edges and onto the backs of tablets.



[Figure 1: 3D Scan of the Obverse and Right Side of a “Micro-tablet” in the Johns Hopkins Archeological Collection. Notice the rounded shape of the tablet, the three dimensional impressions of the wedges in the clay, and the writing streams extending onto the side of the tablet.]

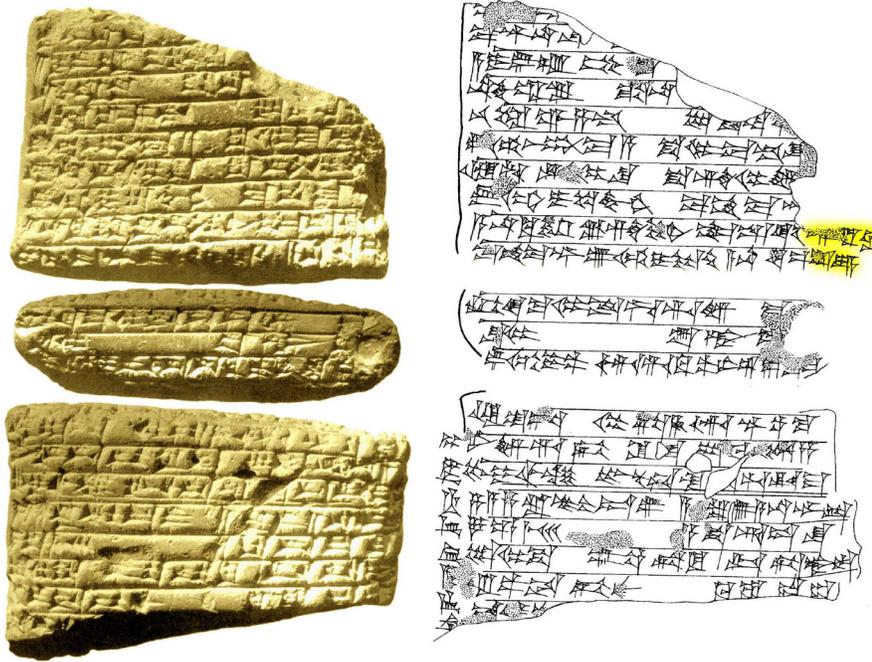
Heretofore cuneiformists have used two main techniques to represent cuneiform documents graphically - 2D photography and hand-drawn copies (autographs).

Photographs have the advantage of conveying an accurate 2D visual representation of a tablet with its text. But due to the multi-tiered three dimensionality of cuneiform documents (wedge impressions, round tablets, and over-the-edge writing), scores of photographs are required, taken at different angles, with different lighting, and at different magnifications, in order to convey enough useful 2D information to enable the collation of a single tablet. This is an onerous burden for research and publication. And even then, those scores of photographs will not contain the amount of information in a single, high resolution 3D scan.

The primary advantage of an autograph is also its primary disadvantage - it provides a representation of an author’s interpretation of what signs are on a tablet. Manually drawing autographs is a laborious, time consuming, error-prone, and highly subjective process requiring direct access to the tablets. Cuneiformists must apply for travel grants to visit the tablet collections in London, Philadelphia, Aleppo, etc. And, though the resulting autographs have the advantage of recording a scholar’s interpretation of difficult to read signs, both the quality of the interpretation and the quality of the drawing vary widely, and disputed readings are common. And

in order to verify disputed readings, cuneiformists must apply for additional travel grants to inspect the tablets once again. The entire process is obviously slow, delicate, expensive, tedious, and, in the end, unproductive. Moreover, unlike photographs, autographs, by definition, are useless for collation, and are practically useless for paleography.

In order to compensate for the deficiencies inherent in both techniques, lately we are seeing more and more cuneiform publications using both 2D photographs and autographs. But the combination is still inadequate for tablet collation. It is no wonder then that we are also seeing a spate of recent forays into 3D surface scanning of cuneiform tablets.



[Figure 2: Photograph and Autograph of an Old Babylonian Text.

The signs highlighted in yellow on the hand copy run off the right edge of the tablet and are not visible in the 2D photograph. “Sargon, the Lion”, Joan Goodnick Westenholz, 1997, *Legends of the Kings of Akkade*]

State of the Art in Cuneiform 3D Surface Scanning

We mention here only a sample of the more interesting, from our point of view, 3D technologies currently being applied to cuneiform research.

Image-based Re-lighting

<<http://www.hpl.hp.com/news/2000/oct-dec/3dimaging.html>>

Tom Malzbender, of Hewlett-Packard, in cooperation with Bruce Zuckerman of the University of Southern California and Walter Bodine of Yale University, has developed a novel approach to digitizing cuneiform tablets. His image-based re-lighting technology yields beautiful images, available for some interesting digital manipulations. But with this technique no tablet geometry is captured, and therefore no 3D manipulation of the tablets can be done. This technique may however prove to be useful when combined with 3D scanning.

Cuneiform Visualization

<<http://graphics.stanford.edu/~seander/cuneiform/>>

Marc Levoy and Sean Anderson of Stanford University have developed a technique for “unwrapping” 3D cuneiform tablet scans and rendering them in two dimensions. This has great potential for cuneiform print publication. They use 3D data acquired by the NRC technology described next.



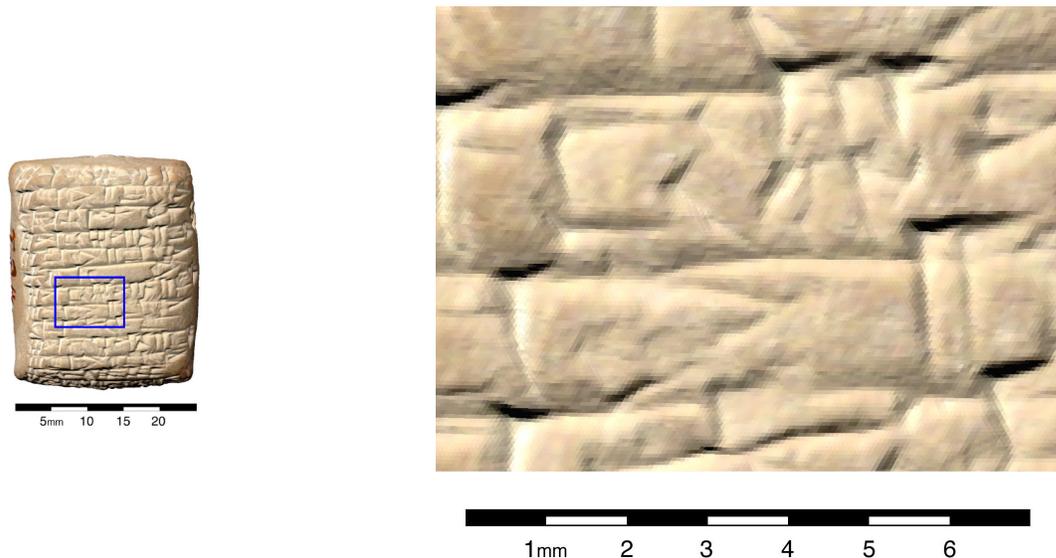
[Figure 3: Unwrapping a Cuneiform Tablet]

Tri-Color Laser Triangulation

<<http://www.vit.iit.nrc.ca/VIT.html>>

Marc Rioux and the team at the Visual Information Technology Group of the National Research Council of Canada hold several patents on tri-color (red, green, and blue) laser triangulation surface scanning. This ingenious method simultaneously captures, at 50 micrometer resolution, both 3D geometry and colorimetric information. We believe this technology is the candidate of choice for further development for our application.

After reviewing 3D technologies for the past 2 years, it is our assessment, and that of others, that all current 3D technologies are inadequate for tablet collation; they are too slow, too expensive, and, most significantly, achieve too low of a resolution.



[Figures 4 & 5: 3D Scan of the Obverse of a “Micro-tablet” in the Johns Hopkins Archeological Collection. In some ways this tiny tablet presents a worst-case scenario for resolution testing. This enlargement from the highlighted area of the tablet gives some indication of the current resolutions attainable with laser triangulation - ca. 50 micrometers, or 500 dpi. We estimate our resolution requirement to be approximately 25 micrometers, or 1000 dpi.]

Digital Hammurabi Project

The mission of the Digital Hammurabi team is to make available to every researcher’s computer very high resolution, three dimensional models of cuneiform tablets along with the computer encoding needed to type those texts in.

Building on an initial 3-year, \$1.6 million grant from the U.S. National Science Foundation, we are addressing all issues related to achieving that mission. The specific goals of the project are to:

- 1) design and produce a portable, non-contact, user-friendly, very high resolution 3D surface scanner that can scan all facets of an average cuneiform tablet in minutes while implementing scan-time adaptive resolution down to 25 micrometers, i.e., 40 dots per millimeter (or 1000 dpi);
- 2) develop new computer algorithms for stitching gigabytes of raw 3D data together into coherent, virtual tablets for real-time manipulation and rendering over the Internet;
- 3) coordinate a formal proposal to the Unicode Consortium for a standard Sumero-Akkadian cuneiform computer encoding.
- 4) collaborate in the development of international XML standards for 3D data, cuneiform bibliography, and cuneiform text markup, aimed at feature comprehensiveness, data longevity, and data integrity;
- 5) establish a petabyte-scale digital library and museum of virtual 3D cuneiform tablets targeted for rapid, real-time Internet2 dissemination;
- 6) invent a completely new technology - automated 3D character recognition of cuneiform writing.

Why Digitize?

There are manifold benefits to digitizing cuneiform tablets:

- * High quality digital surrogates of the tablets can serve to archive these unique cultural artifacts.
- * Virtual tablets can help with re-patriation issues, particularly in serving as high quality substitutes for the originals.
- * Distributed, archival storage of digital surrogates provides an increased level of disaster-proofing.
- * There is a lessened need for researchers to handle original artifacts, thus improving the chances for preservation.
- * Digitization allows for accurate and inexpensive 3D reproduction of the originals in various materials, including plastic, paper, stone, metal. (3D printing, <<http://www.cfoeurope.com/ecfo0701d.html>>, or rapid prototyping, <http://home.att.net/~castleisland/web_lks.htm>)
- * Students will have vastly superior tools for study.
- * Researchers will be less dependent upon expensive travel.
- * 3D models will enable geometrically-accurate 2D representations of the objects.
- * Virtual tablets, once digitized, are not subject to environmental degradation, thereby enhancing preservation of information.
- * Researchers will be able to programmatically search for patterns in digitized 3D data. For example, they can perform 3D optical character recognition on digitized tablets.
- * Digitization opens up broader access to items typically stored out of sight.
- * Digital publication will increase interest in the originals.

The greatest drawback to digitized information, and one we will have to address, is its vulnerability to fraud. Verification of authenticity is one of the most serious issues facing purveyors of digital information. (In our view data longevity concerns are pretty much taken care of by XML.)

Current Status of the Project

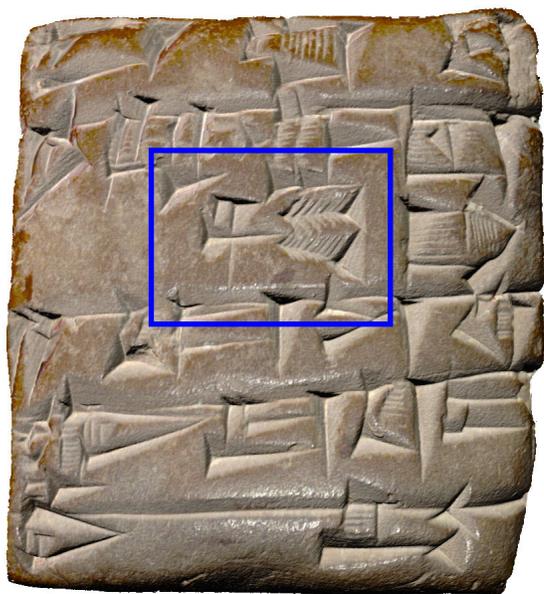
At the end of the first four months of our three year NSF grant, we can report significant developments in the project.

* Culminating a two-year process of 3D technology review, the Digital Hammurabi team has provisionally **settled on the tri-color laser rangefinding technology** invented at the National Research Council of Canada as the technology of choice for further development in 3D surface scanning of cuneiform tablets. In order to meet our requirements we foresee the need to advance the state of the art in the following ways:

- higher resolution
- faster data capture

- smaller apparatus size
- vibration control or cancellation
- automated acquisition of next-best-views

After viewing surface scans of cuneiform tablets on 3D displays and looking at the NRC 3D datasets for some of the tablets they scanned for us, we have relaxed our initial requirement of 10 micrometer resolution to 25 micrometer resolution. We feel that 3D displays increase the information content of the data two- or three-fold in our application. And 3D display technology is poised for rapid expansion into the consumer market space over the next two or three years. (<<http://www.3dcgi.com/index.htm>>)



[Figures 6 & 7: 3D Scan of a Small, Dark, and Somewhat Shiny Cuneiform Tablet in the Hopkins Archeological Collection. Enlargement of One Cuneiform Sign. This size of this sign is typical for a large number of cuneiform texts. And the 50 micrometer/500dpi resolution is almost good enough for tablet collation. We are planning to double the resolution to 25 micrometer/1000dpi.]

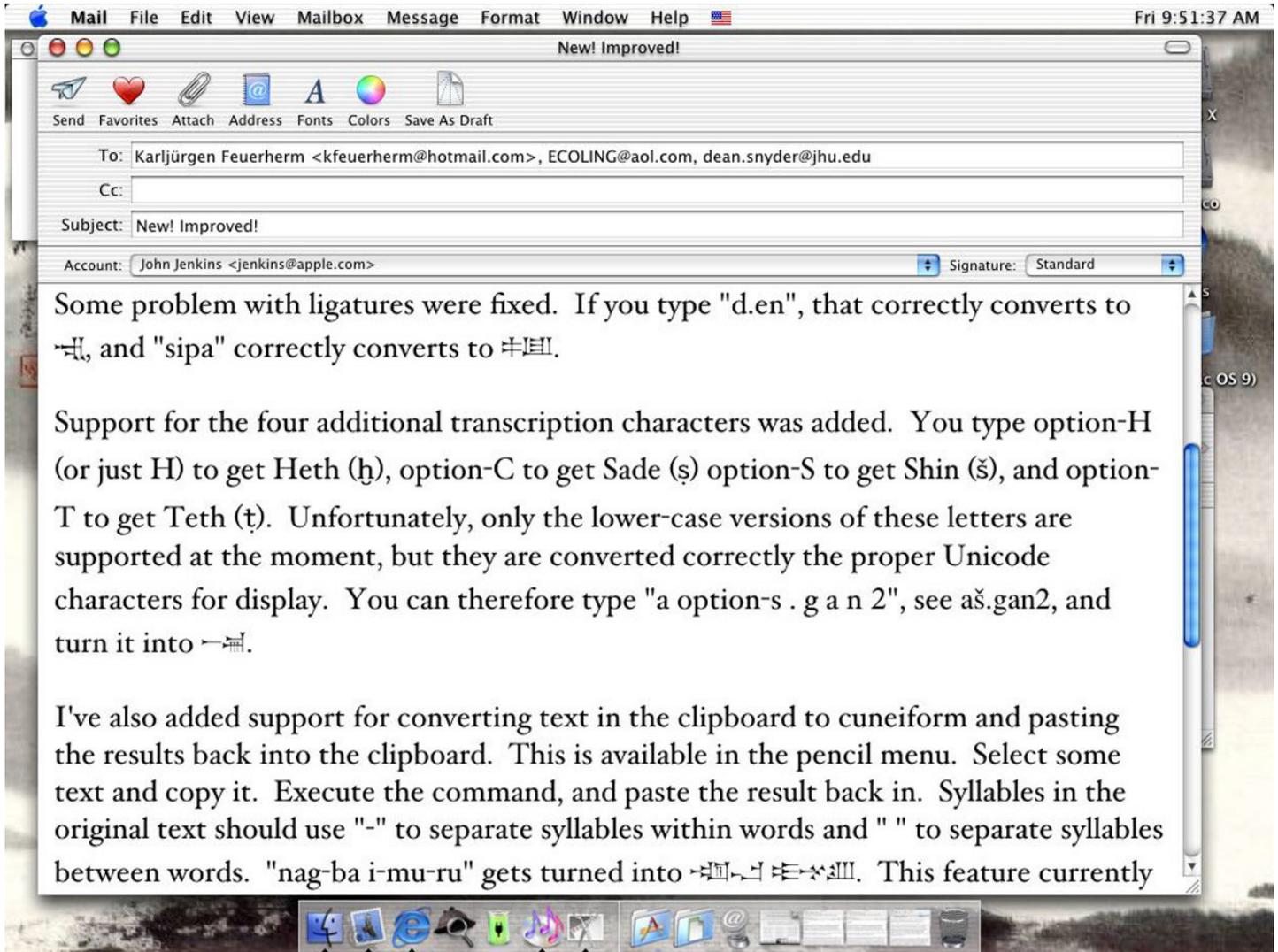
As of this writing we are **in active discussions with the Visual Information Technology Group of the National Research Council of Canada** about forming a research collaboration centered around enhancing their tri-color laser rangefinding technology. Both parties are anxious to make this collaboration happen, but it is too soon to announce anything.

* **Dr. Rykle Borger** has graciously made available to us materials from the forthcoming 3rd edition of his discipline-standard, cuneiform sign list, Mesopotamisches Zeichenlexikon. We are using his materials as a point of departure for discussions on what to include in the Unicode proposal. This database of cuneiform signs is on the web at <<http://www.jhu.edu/ice/signs/>>.

We are **sponsoring and hosting the second ICE conference** at Johns Hopkins University in the Spring of 2003. (<<http://www.jhu.edu/ice/>>)

In November 2003, we are **hosting the annual Unicode Technical Committee and L2 meeting** at Hopkins, at which time we intend to formally present our proposal for cuneiform encoding. (<<http://unicode.org/consortium/utc.html>> & <http://www.ncits.org/tc_home/l2.htm>)

Soon we will be able to send plain text cuneiform email to one another!



[Figure 8: The World's First Cuneiform Email.

Using a prototype Unicode font and input method in Mac OS X, Apple System Software Engineer John Jenkins sent the world's first cuneiform email to other members of the Initiative for Cuneiform Encoding working group on March 30, 2001.]

* We continue to **collaborate with the Cuneiform Digital Library Initiative** at UCLA (<<http://cdli.ucla.edu/>>), as we work towards enabling the interlinking of our digital assets. We were actively involved in the writing of their copyright agreement used with curatorial staff of museums in preparation for their internet distribution.