Digital Hammurabi

48th Rencontre Assyriologique Internationale, 2002
Leiden, Netherlands

There are at least two major technological obstacles to more productive cuneiform research:

* the current inadequate state of the graphic representation of cuneiform tablets
* and the lack of a standard computer encoding for cuneiform text

Graphic Representation of Cuneiform

Cuneiform is typically, 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 on wet clay

* the writing streams are three dimensional - sentences often run over the edges and onto the backs of tablets

Two dimensional photography is inherently inadequate for rendering the multi-tiered three dimensionality of cuneiform, which is why we are seeing a recent flurry of activity in the 3D surface scanning arena.

State of the Art in 3D Surface Scanning

I present here only the most interesting, from our point of view, 3D technologies.

Image-based Re-lighting
Tom Malzbender, Hewlett-Packard (in cooperation with Bruce Zuckerman, University of Southern California and Walter Bodine, Yale University)
<http://www.hpl.hp.com/news/2000/oct-dec/3dimaging.html
This novel technology yields beautiful images, but no tablet geometry is captured, and therefore there can be no 3D manipulation of tablets.

Laser Triangulation
Arius3D, Toronto, Canada
<http://www.arius3d.com>
This is probably the best current implementation of the candidate of choice for the surface scanning of cuneiform tablets - laser triangulation. The Arius3D method (using ingenious patented technology invented by Marc Rioux of the Canadian Research Council) captures, at the same time, both 3D geometry and colorimetric information. The problems with this implementation of the technology are its enormous expense and its insufficient (100 μicron) resolution.

Cuneiform Visualization
Marc Levoy & Sean Anderson, Stanford University
<http://graphics.stanford.edu/~seander/cuneiform/>
Levoy and Anderson's technique for "unwrapping" a 3D cuneiform tablet scan onto a 2D surface has great

potential for cuneiform print publication. (They use the Arius3D technology to acquire their 3D data.)

In short, all current 3D scanning technologies are inadeqate for tablet collation, being, as they are, too slow, too expensive, and of too low a resolution.

For these reasons cuneiformists continue to depend, in their research, publications, and teaching, primarily upon hand drawn copies, occasionally supplemented with photographs. But producing 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 cuneiform autographs are practically useless for certain modes of research, such as paleography.

Computer Encoding of Cuneiform

Cuneiform research is even further 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 sophisticated and complicated text analysis software and fonts needed for greater productivity in cuneiform philology.

Digital Hammurabi Project

The Digital Hammurabi Project addresses both of these issues.

Digital Hammurabi is a major, cross-discipline effort originating at Johns Hopkins University aimed at making very high resolution, three dimensional models of cuneiform tablets available to every researcher's computer and at producing an international standard Unicode encoding for cuneiform text.

When successfully completed, the project will enable scholars to select tablets from cuneiform digital libraries for use on their local computers where they can manipulate them at will, linking graphic cuneiform to encoded cuneiform. Scholars will be able to pan, tilt, rotate, magnify, and re-light these virtual tablets. They can produce "unwrapped" two dimensional projections of 3D tablets for print. They can generate accurate 3D plastic models of tablets. They can apply sophisticated and powerful text and corpora processing software toward concordance generation, morphological analysis, proximity and contextual searching, and automatic generation of critical apparatuses. Automated 3D character recognition will become a reality.

The Digital Hammurabi Project has just been awarded $1.6 million over the next three years by the U.S. National Science Foundation. This grant provides us with a good start toward achieving our goals.

Specific Goals of the Project

The Digital Hammurabi Project plans to:

1) produce a portable, non-contact, user-friendly, very high resolution 3D surface scanner that can scan all facets of an average cuneiform tablet in under a minute while implementing scantime adaptive resolution down to 10 micrometers (i.e., 100 lines per millimeter - at least 4 times finer than currently available resolutions) [Although there will always be a need to personally inspect tablets for the more difficult readings, we expect high quality

3D renderings of cuneiform tablets to obviate the need for personal inspection in approximatley 90% of the cases.]

2) develop new computer algorthims to stitch gigabytes of raw data together into coherent, virtual tablets for real-time, multi-resolution rendering, self-shading, and manipulation by researchers over the Internet using software of our own design

3) coordinate a formal proposal to the Unicode Consortium for a standard Sumero-Akkadian cuneiform computer encoding (continuing ICE, the Initiative for Cuneiform Encoding - <http://www.jhu.edu/ice>) The encoding proposal is intended to include characters from Sumerian, Akkadian, Eblaite, Hittite, Elamite, and Hurrian, but not Old Persian. (A somewhat related proposal for alphabetic Ugaritic cuneiform has already been accepted for Unicode consideration.)

4) collaborate in the development of new international standards for 3D data aimed at data longevity and data integrity

5) collaborate in the development of new international standards for cuneiform text markup (XML metadata), aimed at feature comprehensiveness, data longevity, and data integrity

6) establish a leading petabyte-scale digital library archive of virtual 3D cuneiform tablets targeted for rapid, real-time Internet2 dissemination

7) invent a completely new technology - automated 3D character recognition of cuneiform writing.

The Johns Hopkins team is composed of:

* one faculty member from the Department of Near Eastern Studies - Jerrold Cooper

* three scientists from the Johns Hopkins Applied Physics Laboratory (a major laboratory for space and defense research) - Donald Duncan, Brad Boone, and Kevin Baldwin (in physics, optics, and instrumentation)

* two faculty members from the Department of Computer Science - Subodh Kumar and Jonathan Cohen

* and two members from the Sheridan Libraries at Hopkins - Lee Watkins and Dean Snyder

Though the full realization of these goals will take several years, our thrust during the three years of this initial NSF grant is to develop a working high resolution scanner, computer algorithms for multi-resolution rendering and self-shadowing of 3D tablets, and the beginnings of a digital library infrastructure to support an accumulating archive.

In conclusion...

The Digital Hammurabi Project is applying the very latest computer technologies to these oldest of written documents in the hopes of making them more widely available to scholars and more accessible to better tools for philological research. We fully expect the new hardware and software technologies we develop to revolutionize cuneiform studies, not only by enabling plain text cuneiform transmission and analysis and by providing for ubiquitous 3D access to the world's tablet collections, but also by limiting physical contact with these valuable and unique ancient artifacts, while at the same time preserving our heritage through redundant archival copies of the originals, thereby ensuring their preservation into the future.

The success of the 3D portion of the project hinges on advances in hardware and software technology, which will generate several doctoral dissertations, research papers, and international workshops. The technological fallout is expected to enrich other disciplines.

The success of the encoding portion of the project depends upon the active involvement of a broad spectrum of cuneiform scholars, specialists in the various languages, genres, and areas. If you haven't already, please join the ICE email discussion list by giving me or Jerry Cooper your email address. This is where the main discussion takes place on issues surrounding the encoding of cuneiform. In addition, we will be posting shortly to the ICE website, <http://www.jhu.edu/ice>, a Java and XML-based database of cuneiform signs used to keep track of the current official state of the Unicode proposal. (Dr. Rykle Borger has kindly provided to the Initiative for Cuneiform Encoding pre-publication access to his forthcoming Mesopotamische Zeichenliste materials. We are using his materials as a starting point for discussion on character repertoire and ordering.) The main issues to be settled before we can submit an offical and complete encoding proposal to the Unicode Consortium are:

1) Exactly what characters will be included in the proposal? (finalizing such issues as the sign inventory and the treatment of "punctuation", numerics, textual adjuncts, sign mergers and splits, compound and complex signs, etc.)
2) What will the character names be?
3) What order will we impose on the characters?

The Digital Hammurabi team is actively seeking open collaboration with scholars and curators everywhere in every phase of our project. As we develop new hardware and software solutions we plan to make them broadly available.

If you are interested in becoming involved please contact Jerrold Cooper (<anzu@jhu.edu>) or Dean Snyder (<dean.snyder@jhu.edu>) and view our soon-to-be-published website at <http://www.jhu.edu/digitalhammurabi>.