Unicode fonts for Cuneiform
Old Babylonian (crusive and monumental script), Neo-assyrian, Hittite

I. Unicode Standard 5.0 and Cuneiform: general principles and critical notes
II. Cuneiform fonts: Old babylonian, Neo-Assyrian, Hittite
>   *Installation*
>   *Sign List and notes*
>   *Automatic text*

I. UNICODE STANDARD 5.0 AND CUNEIFORM

The *Unicode Standard version 5.0* (July 2006, see http://unicode.org) now includes blocks for *Cuneiform* (12000-1236E) and *Cuneiform Numbers and Punctuation* (12400-12462 and 12470-12473).

Unicode is an computer standard which assign to each character a unique number (code point), wathever the operating system, the software and the language may be. This standard was developped during the times for many writing systems, but was lacking for Cuneiform script. *Initiative for Cuneiform Encoding* (ICE http://www.jhu.edu/ice/), founded in Baltimore in 2000, aimed to fill this gap. In 2006, a final proposal for the cuneiform writing system was elaborated by Steve Tinney, Michael Everson and Karljürgen Feuerherm (http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2786.pdf).

The assignment of new blocks for Cuneiform has the purpose to give a solution that non-Unicode fonts could not provide. Limited to 256 characters, they do not permit the encoding of signs in a single font. Two, three, and even four fonts were necessary to have the whole set of signs for Old babylonian, Hittite or Neo-Assyrian. The encoding itself was different from one font to antoher. In this point of view, the Unicode proposal for Cuneiform would make easier the encoding and the use of the fonts. On the whole, the results are more than positive and we immediately see the advantages: electronical corpora, lexicographic databases, research on corpus (sorting, occurrences, index... see http://www.jhu.edu/ice/)..

As every innovation, the Unicode Standard 5.0, even if it meets the major needs, gives rise

to some difficulties and problems, which probably would need some improvements. The main critical review has been done by R. Borger in the introduction to *List of Neo-Assyrian Cuneiform Signs. A practical and critical guide to the Unicode blocks «Cuneiform» and «Cuneiform Numbers» of Unicode Standard Version 5.0* (http://www.sumerisches-glossar.de/download/SignListNeoAssyrian.pdf), compiled by M. Studt (D. Bachmann's revised list). See also the article Unicode cuneiform (2007, August 24): *Wikipedia, The Free Encyclopedia* (http://en.wikipedia.org/wiki/Unicode_cuneiform).

The final proposal sets out the general principles for encoding.

1) The character inventory is based on the Ur III sign list, compiled by the *Cuneiform Digital Library Initiative.* As said, this is a first stage in the definition of code points; other stages should concern other periods (Old Akkadien, 2334-2154 et Early Dynastic, 2900-2335, Archaic Cuneiform).

As Borger wrote, the proposed list is limited, in the sense that it does not account for different periods and regions using cuneiform writing. The consequence is that some signs used by Hittite or even Neo-Assyrian do not have code point for Hittite or even for Neo-Assyrian (see the lists). Several code points should be added to get a complete list, concerning all chronological and geographical aspects of cuneiform writing system.

2) Cuneiform sign and cuneiform characters do not necessarily correspond; complex signs and compound signs are distinguished (see final proposal, p. 7).
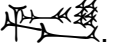
°*complex* signs are made up of primary sign with one or more secondary sign written within it or otherwise conjoined to it. The whole is a unit. For example, the single sign KA is used to form complex sign KAxGU (where GU is written in KA), KAxLI ... The complex sign being a unit, it has a specific code point, in this example, KA = U12157, KAxGU = U1216B, KAxLI = U12172.

Some signs of Borger's List in *MeZL* do not have code point. These are rare signs, KAxTU, KAxḪAR, LAGABxGI etc., about thirty, fourty signs. Others are attested for Hittite: KAxÀŠ, KAxÚR, KAxGAG, KAxGIŠ, KAxPA, KAxLUM, SIxSÁ, EZENxŠE, AMARxKU$_6$, ÁBxA, KISIM$_5$xÚ-MAŠ.

° *compound* signs are made of two or more signs organized in linear sequences. Each sign exists in other respects as single character. The whole is generally viewed as a unit, but each component will be separately encoded. For example: IDIGNA, composed of MAŠ +GÚ+GÀR, code points U12226+U12118+U120FC .

The *Unicode Standard 5.0* list of Cuneiform signs is arranged according to the latin alphabet and gives an "etymological" description of the signs (simple, complex and compound, see Cuneiform Unicode, Wikipedia).

This list, Borger says, "introduces new set on conventional readings, without offering any explanation", instead of known and generally used conventional readings in Assyriology. The aim is to describe the sign (Cuneiform sign A, Cuneiform sign KA times LI, etc.), but the use often is not practical (we have to know that GIR = ḪAgûnu, AMAŠ = DAG KISIM5 times LU + plus MASH2). The most critically point of view about this new way deals with the sign analysis on the one hand, and the sometimes laborious use of code points. The "splittability" of some signs actually seems arbitrary and could be discussed ; see for example Borger, *MeZL*, n° 754 MEŠ, which has to be decomposed ME+U+U+U (ME = U12228; U+U+U = U1230D) or n° 459 DUL splitted U+TÚG (U1230B+U12306) according to the *Unicode Standard*; but according to Borger, these signs are not really "splittable". In other cases, it will be necessary to combine one, two, or three code points: U12226+U12118+U120FC for IDIGNA, U12263+U121EC for TÙR (NUN-LAGAR), etc. Ligatures are in theory impossible, unless another font is made, which contains the needed variants (that could be absolutely artificial, see, for Hittite, notes with the sign lists). For Old babylonian and Neo-Assyrian sign NIGIN (*MeZL*, n°804) = LAGAB-LAGAB (U121B8+ U121B8), we get ⸢cuneiform⸣, without the possibility of a real ligature.

In other cases, the split of signs is difficult, like for Hittite, which uses signs based on an Old Babylonian cursive from North Syria (E. Neu - C. Rüster, *Hethitisches Zeichenlexikon*, Wiesbaden, 1989). The sign IDIGNA, as used in Hittite syllabary, is not splittable in practice: U12226+U12118+U120FC gives ⸢cuneiform⸣, not very like the "real" sign ⸢cuneiform⸣. For this kind of compound sign, we have made the resolution to isolate artificially each component, that means to design variant for MAŠ, for GÚ, for GÀR, which permit to get the final sign. This kind of variant is pointed out in the footnotes in the sign list. The use of automatical text will also permit to get round this problem (see below).

II. CUNEIFORM FONTS: OLD BABYLONIAN, NEO-ASSYRIAN, HITTITE

The Unicode Cuneiform fonts (TTF) have been designed by Sylvie Vanséveren.

- Santakku: Old Babylonian, cursive script
- SantakkuM: Old Babylonian, momumental script (Hammurabi)
- Assurbanipal: Neo-Assyrian
- Ullikummi (A, B, C): Hittite.

They are freely available for the scientific community. They can not be altered or solded.

Fonts installation

°MacOSX: put the fonts into User/Library/Fonts.

°Windows (XP):

      click on Start, then Configure settings

      click on Fonts: the font folder opens

      two ways for adding fonts in Windows:

        - direct copy of fonts in the font folder

        - click on File menu in the font window and choose Install new font. Select the font in the List of fonts. Click on OK.

Sign List and notes

Each sign list contains the following:

      - MeZL numbering

      - Unicode code point

      - sign (conventional reading and Unicode description)

      - sign in the font (and variant for Hittite)

      - autotext

Explanation notes and remarks in each list concern some specific problem about a sign, about the split of signs, and about the necessary variants for composing complex and compound signs (almost for Hittite).

Automatic Text

Once the encoding is made, it remains to use the fonts in concrete terms, that is to write in cuneiform. It always is possible to access to the signs via the Character Palette in Mac OSX or via the Special Character menu in text editor, but this method is laborious and tedious, since there is necessary to know the signs, or recognize them, or to know the code point of each specific sign. In order to make things easier, one can use automatic text.

Automatic text insertion is possible in text editors like Word and Nisus Writer Pro. File templates are given here for Word and Nisus. **Each template can be modified by the user, who can add, modify or delete entries.**

NB. Neo- et OpenOffice (Mac et PC) do not seem to correctly embed the fonts (nor via

special character insertion, neither via Character Palette in OSX). Other text editor (like Mellel for Mac) do not offer autotext.

The principle consists in associating an automatic text entry with a sign, or group of signs. It is so possible to write in Cuneiform without having to know the code point of the signs. In order to do that, some latin characters are encoded in the fonts (characters from the Cuneitruetype font with kind permission of Prof. Dominique Charpin).

So typing for example the entry "ma", and then select automatic insertion (via Insertion menu or keyboard shortcut, see below), we obtain the sign (U12220)

      𒈠 (Assurbanipal)

      𒈠 (Santakku)

      𒈠 (SantakkuM)

      𒈠 (UllikummiA)

In the same way, the entry "lugal" will give (U12217)

      𒈗 (Assurbanipal)

      𒈗 (Santakku)

      𒈗 (SantakkuM)

      𒈗 (UllikummiA)

For compound signs, made of two or more signs in linear sequence, for each of which one code point is assigned, the principle permits to associate the signs:

"mac" (MAŠ)	U12226	𒈦	𒈦
"gu2" (GÚ)	U12118	𒄭	𒄭
"gar3" (GÀR)	U120FC	𒃼	𒃼

"idigna"   𒀀𒄭𒃼 (Assurbanipal)

                 𒀀𒄭𒃼 (SantakkuM)

NB. Only one template is proposed here for Old Babylonian and Neo-Assyrian (Cuneiform.dot); the automatic entries are the same. Since Hittite regularly has other values for the signs, another template (Hittite.dot) contains the insertions based on the Hittite values.

Variants are marked with "v" or "vv" after the entry

Assurbanipal: me3   me3v

Ullikummi: mah   mahv   mahvv

MICROSOFT WORD

For more informations, see Word Help, or http://support.microsoft.com/kb/207925/en-us

- **Template**: Cuneiform.dot, Hittite.dot

- **Installation**:

    °MacOSX: Applications\folder Microsoft Office \Templates (or in the folder My templates)

    °Windows (XP): C:\Windows\Documents and Settings\User\Application Data\Microsoft\Templates (hidden folder by default).

- **Choose template** via Library projects.

Use of autotext

There is two different ways to proceed:

- Insertion menu → Automatic Insertion. A rolled list gives the different insertions.

- Keyboard shortcut for automatic insertion.

To access to the dialog box for defining entry of automatic text:

- Insertion menu → automatic insertion (top of the rolled list)

- or click on button "automatic insertion"

- or via Tools menu → automatic correction

To define a keyboard shortcut for the command automatic insertion:

Tools menu → Customize keyboard

In the dialog box, select on the left All the commands

Select AutomaticInsertion and choose a keyboard shortcut.

To access to the button " automatic insertion "

Tools menu → Customize toolbar

Drag the button on a toolbar (Standard toolbar for example).

NISUS WRITER PRO (MAC only)

In Nisus Writer Pro, automatic text is linked with a glossary. A glossary, when loaded, is available for all documents, whatever template is used. The Cuneiform glossary is a file with cuneiform signs and associated entries. For more informations, see the NWP guide.

Format: Cuneiform.ngloss

Installation:

> User/Library/Application Support/Nisus/Glossaries

> the glossary can be imported via Preferences → Quickfix → Import glossary (choose Cuneiform.ngloss).

Use of glossary

> enter the text entry

> entries are expand in Edit menu → expand glossary entries

> glossary entries can also be expanded manually with a combination key (⌘D)

> you can also choose **"Automatic expansion of glossary entries as you type"** in Preferences → Quickfix

To add, delete, or modify glossary entries, open Cuneiform.ngloss and proceed. Each entry must be separated by a Glossary entry break (Insert menu). Save the file.

Sylvie Vanséveren

Research Associate F.R.S.-F.N.R.S.

Université Libre de Bruxelles

svsevere@ulb.ac.be