

Characters Map for Cuneiform Writing System

خارطة الرموز الخاصة بأنظمة الكتابة المسمارية

¹Dhamyaa A. AL-Nasrawi, ¹Huda F. AL-Shahad, ¹Inas R. Shareef

¹Computer Science Department, College of Science,
University of Kerbala, Karbala 56001, Iraq
E-mail: dhmyaa@gmail.com

¹ Computer Science Department, College of Science,
University of Kerbala, Karbala 56001, Iraq
E-mail: huda_msc2006@yahoo.com

¹ Computer Science Department, College of Science,
University of Kerbala, Karbala 56001, Iraq
E-mail: Inas_shareef@yahoo.com

Abstract

Writing system method for a situated about unmistakable marks that would related, by convention, on exactly specific structural level about language. Cuneiform speaks to a standout amongst those most punctual What's more practically persuasive Writing systems of the reality. The aim of this paper is to develop the character map utility included with Microsoft Windows operating systems to view the characters in designed font for special ancient Cuneiform writing system using Unicode standard. The proposed method focused on replacing the block of Cuneiform symbols in Unicode which located in Plane1 by the block in Plane0, in another word, change the glyphs of symbols in Plane0 to desired glyphs. The vitality of the paper one task lies Previously, preserving the history and legacy of ancient Writing systems and make it accessible as naturally manner. It Additionally facilitates should the individuals who interested writers in the ancient languages.

Keywords: *Writing system, Cuneiform Glyphs, Open type font, Character map.*

الخلاصة :

تعد أنظمة الكتابة هي مجموعة من الرموز والعلامات المرئية والمتعلقة بالهيكل الخاص بلغة معينة. الكتابة المسمارية هي احد أنظمة الكتابة القديمة والمهمة في العالم. الهدف من البحث هو تطوير خارطة الرموز المتضمنة مع نظام التشغيل الوندوز لاطهار الرموز المسمارية بواسطة تصميم خط لهذا الغرض والذي يعتمد على ترميز اليونيكود. الطريقة المقترحة تتضمن ابدال كتلة الرموز الخاصة بالكتابة المسمارية الموجودة في اللوح 1 الى كتلة الرموز الموجودة باللوح 0 ، هذا يعني ابدال شكل الرموز الموجودة في اللوح 0 من اليونيكود باشكال جديدة خاصة بالرموز المسمارية.

تكمّن حيوية هذا البحث هو في الحفاظ على التاريخ والارث الخاص بانظمة الكتابة القديمة وجعلها متوفرة ومرئية تلقائيا في خارطة الرموز وكذلك في محرر النصوص، بالاضافة الى التسهيلات للمهتمين ببحوث اللغات القديمة .

1. Introduction

The collection of visible, perceptible or physical signs which used to represent elements of language in a regular style called a writing system. In historical periods, writing used as one of the "signs" of civilization by scientists. While it is a fact that writing systems look to improve in agricultural and urban cultures, it is not a requirement for civilization[1].

Writing, is a set of visible and perceptual marks which are relevant to some specific level of language, these set form of human communication . A writing system, denoted to as a "script" or "an orthography", contains of a collection of visible marks, called characters or graphs that are related to some structure in the linguistic system. Once a character represents a meaningful unit, such as a word or a morpheme, the orthography is called a "logographic writing system"; if it denotes a syllable, it is called a "syllabic writing system"; if a segment of a syllable, it is called a "consonantal writing system"; and if a phoneme, it is called an "alphabetic system". see figure(1)[2]:

	linguistic structure	orthographic structure
meaning-based	text	—
	topic	—
	speech act	pictorial signs
	word morpheme	logographic writing
sound-based	syllable	syllabic writing
	segment	consonantal writing
	phoneme	alphabetic writing
	phone	phonetic alphabet
	feature	featural writing system

Figure1. Types of writing systems

When human societies being more complex, those trying to control them found that their memories were overworked. What they needed was an external store device. What they came up with is writing[3].

The process of connecting over writing began in 3100 B.C by the Sumerians. They began sharing information by pictograms, pictures of specific ideas, animals or objects, and then moved to semasiography, a way of ordering the pictures in a specific order to understand the statement and is easier to identify the meaning. [4].

Cuneiform – or 'Heavenly Writing' as this writing system is also called – represents one of the initial and most powerful writing systems of the world. Today, it is generally assumed that it was generated by the Sumerians, but since the most ancient tablets written in so-called "proto-cuneiform" are not yet decoded [5].

2. Character Map

The Character map is a useful utility involved with Windows operating systems. For any installed font, the purpose of Character map is to view and copy the characters in to the clipboard instate of inputting them, It can be opened via the command line or Run Command dialog using the 'charmap' command. Usually it is suitable for arriving special characters.

There are Unicode fonts which the characters can be gathered by their Unicode sub range. Though the Unicode standard extends character field to plane 16 and many fonts, this tool only supports characters on plane 0 (between U+0000 and U+FFFF) [6].

The aim of this paper is to develop the character map utility included with Windows operating systems to make the characters in designed font for special ancient Cuneiform writing system are visible using Unicode standard. The proposed method focused on replacing the block of Cuneiform symbols in Unicode which located in Plane1 by the block in Plain0, in another word, change the glyphs of symbols in Plane0 to desired glyphs. The importance of the paper project lies in preserving the history and legacy of ancient writing system and make it available as automatically way. It also facilitates to those who are interested authors in the ancient civilization. Rest of this paper is organized as follows. Related works explained in section 3, Section 4 introduces details of Unicode standard, , Proposed method for Ancient Cuneiform Writing System is explained in section 5, Finally conclusions are provided in Section 6.

3. Related work

There are many researches concerned to composing arrangement of ancient writing system and designing related font. Subodh K. *et al.* developed and distributed wide advanced libraries of these aged tablets of wet dirt. They also introduce visualization prerequisites and calculations for processing them [7]. George Williams (George Williams, 2003) describe in his paper the automatic creation of accented characters, and how to add advanced features as well as the basic such as ligatures and kerning pairs to a font[8] . Karel Piška (Karel Piška, 2004) in his paper summarizes skills in creation Type 1 Fonts by converting metafont fonts to PostScript fonts. An improvement process is established with public Indic fonts (Devanagari, Malayalam). Also the features, advantages and

disadvantages of various techniques was discussed in this paper[9]. Hilal M. *et al.* proposed new half and half written work supervisor for Cuneiform dialect with three classifications of symbols (assurian, Sumerian, Akkadian) . Any image might be embedded as picture in suitable size as stated by content font size [10]. Karel Piška (Karel Piška, 2008) produce a cuneiform font collection involve(Akkadian, Ugaritic and Old Persian glyph subsets (about 600 signs) two steps. In his paper generate intermediate Type 1 fonts, and then construct OpenType fonts using FontForge[11].

4.Unicode Allocation

“Unicode” refers to the Unicode Standard which defines character encodings for the writing systems of the world. It is an encoding scheme developed by the Unicode Consortium incorporated under the name Unicode, Inc. in 1991 and the ISO[12]. Any characters in computer need to be coded as numbers. A classic agreement is to use numbers from 0 to 255, because that range fits into a basic unit of data storing and transfer, called a (8-bit) byte[12]. But it would be unreasonable, inadequate, and impossible, to define 256-character codes for all the possible language combinations. Unicode is the solution[13]. Unicode's goal is to solve this issue by using more than one byte for each character[14]. The organization of Code points may range from 0 to 0x10FFFF (= 1,114,111) which divide this range into 17 planes, numbered from 0 to 16. Of these 17 planes, only 6 are currently “populated”, see Figure2[15].

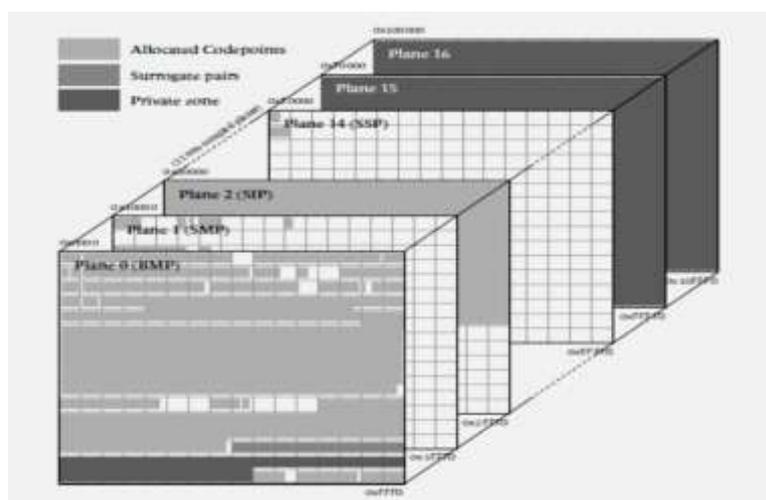


Figure2. Types of writing systems

Basic Multilingual Plane. The Basic Multilingual Plane (BMP, or Plane 0) All textual data can be found in the BMP. It contains the characters for all the modern scripts of the world, also many historical and unusual characters.

Supplementary Multilingual Plane. The Supplementary Multilingual Plane (SMP, or

Plane 1) is devoted to the encoding of characters which either could not be fit into the BMP or see very infrequent usage. This includes many historic scripts, a number of lesser-used contemporary scripts, special-purpose invented scripts, notational systems or large pictographic symbol sets, and occasionally historic extensions of scripts whose core sets are encoded on the BMP.

Supplementary Ideographic Plane. The Supplementary Ideographic Plane (SIP, or Plane 2) is planned as an extra portion area for those CJK characters that could not be fit in the blocks set sideways for more common CJK characters in the BMP. While there are a small number of common-use CJK characters in the SIP .

Supplementary Special-purpose Plane. (SSP, or Plane 14) is the spillover allocation area for format control characters that do not fit into the small allocation areas for format control characters in the BMP.

Private Use Planes. (Planes 15 and 16) are allocated, in their entirety, for private use. Those two planes have a 131,068 characters to supplement the 6,400 private-use characters located in the BMP.

5. Materials and method

In this section, we explain methodology of assignment Cuneiform symbols in characters map, in addition, present font outline and basics of font production.

5.1. Font outline

A font is a photographic communication that is connected to a collection of numbers, characters, and images. A font represents a particular typeface, together with different qualities, for example, size, pitch, and separating. Fonts are utilized to show message on the screen and to print content. Fonts have different styles, for example, italic, strong, and striking italic[16].

A glyphs, are images, drawings, physical marks of black ink on a white background. When the reader enters the digital space for writing, the keys on the keyboard are marked with glyphs; when a key is pressed, a character is transmitted to the system, which, in turn displays glyphs on the screen[17]. The way that the glyph for each character is stored in the individual font-resource file concert to font file data formats. According to this, the font classes are[14]:

- Bitmap fonts: also known as system fonts, raster fonts, , or screen fonts.
- Outline fonts (called vector fonts): a set of lines and curves to define the border of glyphs, which makes the character outlines scalable to any size.
- Stroke fonts: utilize a advance of determined lines and additional data to characterize the profile, or size and state of the line in a particular face which, together, portray the presence of the glyph.

5.2. Font production

The process in which a series of characters typed on the keyboard are transformed into a series of glyphs displayed on the screen is quite complex. It include the encoding of characters and their glyphs.

- Technology of character encoding (Unicode): Unicode is, among other things, a huge catalog of characters including all the writing scripts in the world. One of the purposes of Unicode is to provide a unique code for every character [13].
- Glyphs: Glyphs are the graphic signs of a font. They can be linked to a character code or not, they can represent one or more characters [13].

5.3. Basics methodology of Cuneiform characters map

In this section, we describe the proposed work in detail; Many steps must be followed to assignment Cuneiform symbols in character map. The steps of works was explained in the following:

1. Determine and create the glyphs of Cuneiform symbols.
2. Determine the desired block in Plane0 to replace Cuneiform symbols glyphs.
3. Add specific Features.
4. Test the resulting font.

5.3.1. Determine and create the glyphs of Cuneiform symbols

The word "glyph" plays a major role in a font and a typeface where a glyph can be defined as an element of writing and is a vague term in typography especially in multi lingual fonts. The glyph of letters in electronic font can be created by using (Font Creator) program, This step (design of font) involve all glyphs of Cuneiform symbols. The Cuneiform symbols located in Plane1 started from U12000-U12543. Example of cuneiform glyphs explained in figure 3.

12000		Cuneiform											
	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	120A	120B	120C
0	𐎀	𐎁	𐎂	𐎃	𐎄	𐎅	𐎆	𐎇	𐎈	𐎉	𐎊	𐎋	𐎌
1	𐎍	𐎎	𐎏	𐎐	𐎑	𐎒	𐎓	𐎔	𐎕	𐎖	𐎗	𐎘	𐎙
2	𐎚	𐎛	𐎜	𐎝	𐎞	𐎟	𐎠	𐎡	𐎢	𐎣	𐎤	𐎥	𐎦
3	𐎧	𐎨	𐎩	𐎪	𐎫	𐎬	𐎭	𐎮	𐎯	𐎰	𐎱	𐎲	𐎳

Figure 3 . Glyphs of Cuneiform symbols

This step was covered by take a snapshot for each symbol and save it as .png image to used as glyphs in *font creator* tools, see figure 4 .

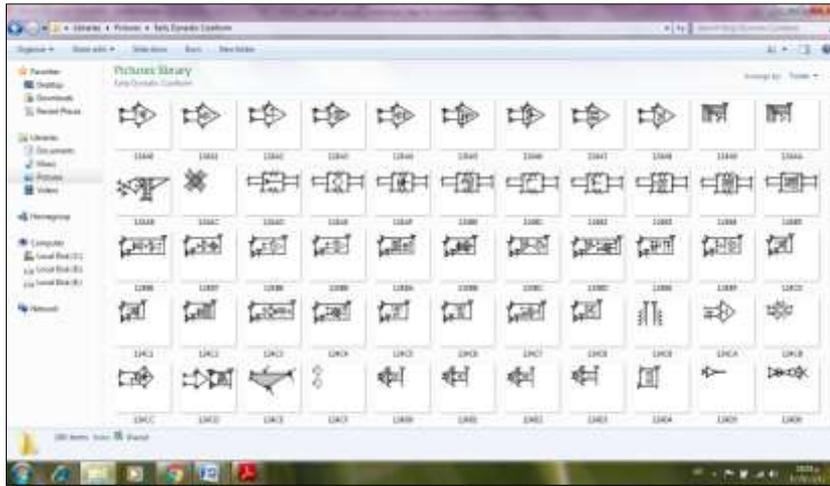


Figure4. Snapshot of symbols

5.3.2. Determine the desired block in Plane0 for replacing

In this paper, we replace the cuneiform symbols located in Plane1 to Plane0 which starting with Latin symbols. Figure 5 shows the Basic Multilingual Plane (BMP) in an expanded format to illustrate the allocation substructure of that plane in more detail, and Plane 1, the Supplementary Multilingual Plane (SMP), in expanded format to illustrate the allocation substructure of that plane in more detail. That mean use the code of Plane0 with glyphs of Plane1.

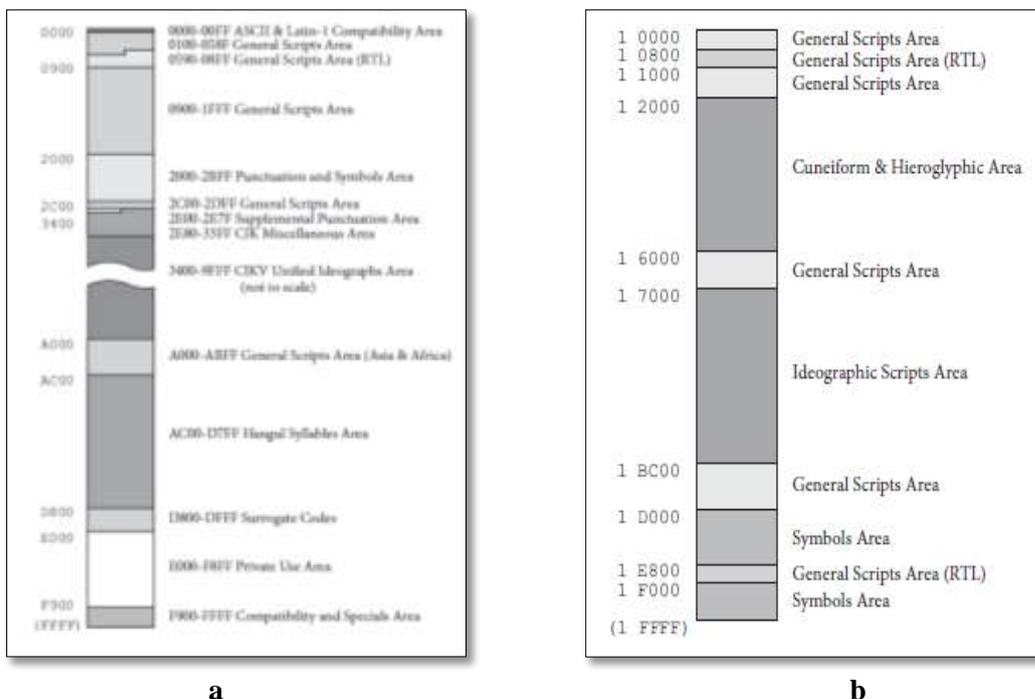
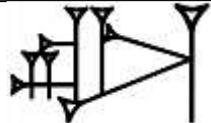


Figure5. a: Plane0 , b: Plane1

Table 1 explain the idea of this work, the glyphs of specific code different from font type to another.

Table 1. Different glyphs for specific code

Font Type	Unicode	Glyphs
Arial	0041	A
Times New Roman	0041	A
Wide Latin	0041	Ⓐ
MV Boli	0041	A
Cuneiform (proposed font)	0041	

5.3.3.Add specific Features

The features of font are kerning between symbols , size of glyphs, define the script (or scripts) that the font will address. The result of this step is true type file with extension .ttf, see Figure 6.

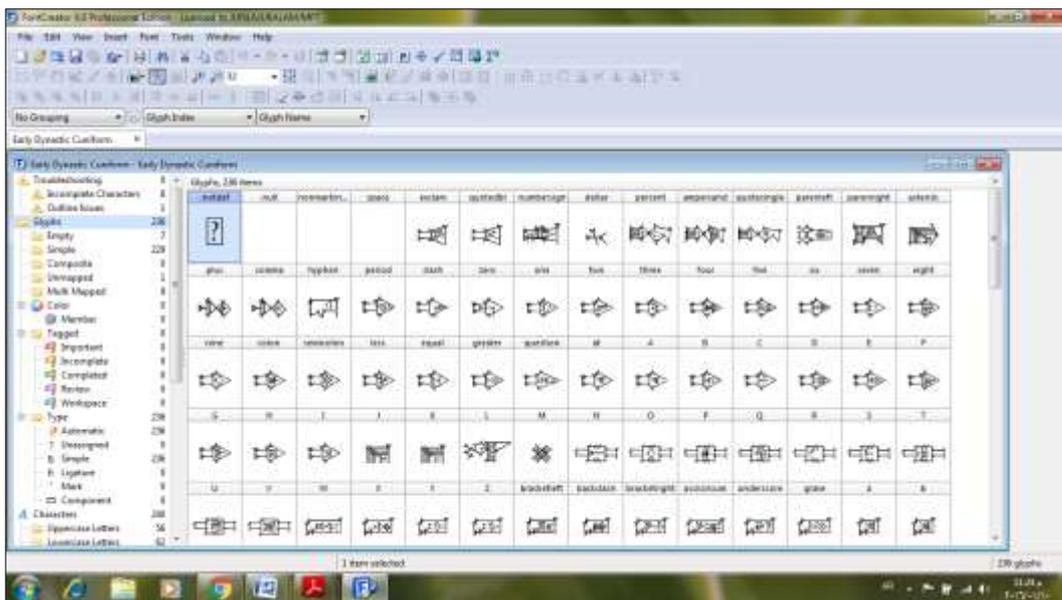


Figure 6. Designed font

5.3.4. Test the resulting font

Finally, when the design and programming of imitated font was done completely, the resulting font must be tested and evaluation. When the evaluation is goodwill, the designed font was installed in computer by name "Cuneiform" and it is already appear automatically in character map to use in any word processing software.

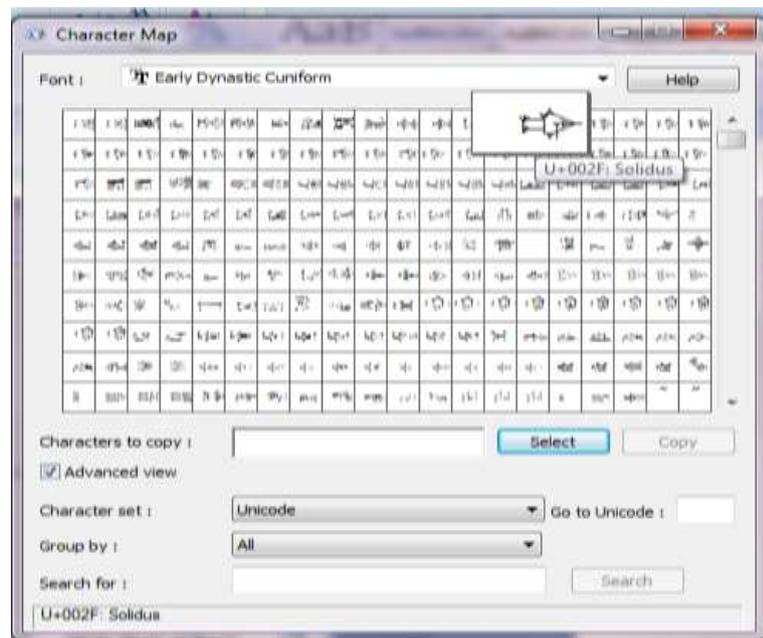


Figure7. Character map of cuneiform symbols

6. Conclusion

In this paper, New methods was proposed to replace blocks in Plane1 to blocks in Plane0. There are no way to display the cuneiform symbols in word processing software , there for the objective of this paper is to appear these symbols automatically in characters map. This works done by font creator tools to design our font. By using Opentype font, the owner can be update the font by add new features to it to increase the number of glyphs in produced font. Finally, this research regard as one a way to preserving the history and legacy of our civilization by involve the symbols in any text editor.

Acknowledgements We thank our colleagues *Dr.Aba Thar AL-Zaidy* who provided insight and expertise as well as comments that greatly assisted and improved this research.

7. References

- [1] Megan W., "Introduction to Arabic Calligraphy", 2009, Available from:
http://campusweb.howardcc.edu/salih/culture/Arabic_Calligraphy.pdf
- [2] Thepublicstudio, "Introduction to Typography", Available from:
<http://thepublicstudio.ca/images/diy/DIY-No5-Typography.pdf>
- [3] Nina K., "Arabic Latin Logo Adaptation", 2010, Available from:
<http://logotalks.com/2010/01/01/arabic-latin-logo-adaptation/>
- [4] George W., "Font Creation with FontForge", TUGboat, Vol. 24 , No. 3 —
Proceedings of EuroTEX 2003.
- [5] Karel P., "Creating Type 1 Fonts from metafont Sources", Lecture Notes in
Computer Science Vol. 3130, pp 240-256, 2004.
- [6] Karel P., "Creating cuneiform fonts with MetaType1 and FontForge",
TUGboat Vol. 29, No. 3, pp. 421-425, 2008.
- [7] Ilham C., Rachid B., Lhadj M., Mustapha H., "Automatic Creation Of
Moroccan Alphabetic", International Journal of Engineering Science and
Technology, Vol. 3, No. 4, pp. 3535-3540, 2011.
- [8] Sherif S., Hossam A, "Experiences with Arabic font development",
TUGboat, Vol. 33, No. 3, pp. 295-298, 2012.
- [9] Sherif S., "Hassan Automating The Generation And Typesetting Of Arabic
Script", Msc. Thesis, Electronics and Communications Engineering
Department Faculty of Engineering, Cairo University, 2015.
- [10] Kazim F., Babak N., "Toward automatic development of handwritten
personal Farsi/Arabic OpenType® fonts", International Journal on Document
Analysis and Recognition, Vol. 18, No. 3, pp. 249-262, 2015.
- [11] Susan M., "100 Things Every Designer Needs to Know About People",
Berkeley,CA, USA: New Riders Publishing. 257 p., 2011.
- [12] Daniel Y. , "Unicode for Under-Resourced Languages", 2006 .Available
from:
<http://mtarchive.info/LREC-2006-Jacob.pdf>
- [13] Jukka K., "Unicode Explained", USA: O'Reilly Media, Inc. 678 p., 2006.
- [14] Yannis H., "Fonts & Encodings", Sebastopol, CA, USA: O'Reilly Media,
Inc. 1040 p., 2007.
- [15] Stephane B., "Arabic Font Production Tutorial Part one", Kahatt Books. 12 p., 2011.