Computerized Hittite Cuneiform Sign Recognition and Knowledge-Based System Application Examples

A. Ziya Aktas, Prof. Dr.,
Beste Yesiltepe, M.S.,
Tunc Asuroglu, M.S.,
Baskent University-Ankara, Turkey

Doi:10.19044/esj.2019.v15n33p32 URL:http://dx.doi.org/10.19044/esj.2019.v15n33p32

Abstract

The Hittites had lived in Anatolia more than 4000 years ago. The Hittite language is one of the oldest and may be the only one still readable and grammar rules are known member of Indo-European language family. The Hittites had a cuneiform script of their own written on soft clay pads or tablets. Tablets made durable and permanent by baking them after writing with some tools. That is why they could endure for thousands of years buried in the ground. The study of Hittite language has been made manually on the Hittite cuneiform tablets. Unfortunately, field scientists have read and translated only a relatively small number of unearthed tablets. Many more tablets are still waiting under and over ground in Anatolia for reading and translation into various languages. To read and translate the cuneiform signs, using computer-aided techniques would be a significant contribution not only to Anatolian and Turkish but also to human history. In this paper, recognition of Hittite cuneiform signs by using computer based image-processing techniques is reported. Additionally, uses of data-mining applications are also included in the paper. Most importantly, the authors also demonstrated feasibility of an expert system on the Hittite cuneiform scripts.

Keywords: Cuneiform sign recognition, Data-mining, Expert System, Hittite cuneiform script, Image Processing and Computer Vision, Optical Character Recognition

Introduction

In Anatolia-Turkey, the kingdom and empire of the Hittites or Hattis as named in the Bible, had ruled nearly half a millennium during the years BC 1650-1200. They were one of the greatest world powers of their time. Hittite language that the Hittites used is one of the oldest members of the Indo-European language family that is still readable and it has known grammar rules. Because of this property, Hittites and Hittite language have become interesting and historically valuable in Western countries like USA, Germany and England, including some others.

As noted by Karasu (2013), Czech scientist Bedrich Hrozny revealed grammar rules of Hittite language in the beginning of the 20th century (in 1915). Since then, reading, translating and interpreting of Hittite cuneiform scripts have needed human manual efforts. In order to read cuneiform scripts and to do necessary translations, we have required expert scientists, who are unfortunately a few globally.

This paper includes a summary of computerized works performed relatively recently in the Computer Engineering Department of Baskent University-Ankara in three consecutive M.S. Theses, namely Dik (2014), Asuroglu (2015) and Yesiltepe (2015). They could help computer-based translation of signs in Hittite cuneiform tablets to Latin script. Hittite cuneiform signs in tablets were read by using some computer-based image-processing techniques and were matched with signs that were already stored in databases and later translated into Latin script. During these studies sign matching performances of the techniques that were used in reading Hittite cuneiform signs were compared. Some techniques to speed up the matching process of cuneiform signs during the study were also proposed.

In Data-mining part of the studies, categorization of Hittite cuneiform signs based on their geometrical features were carried out to speed up the process of reading cuneiform signs in tablets by categorizing similar signs. After categorization of cuneiform signs, data-mining classification algorithms were applied. Comparative classification performances of applied algorithms were reported in the paper.

The major contribution of this paper is to demonstrate the applicability or the technical feasibility of using image processing and computer vision techniques and Knowledge-Based Systems or Expert Systems on the translation of Hittite cuneiform scripts written on clay tablets or their copies. Paper finishes after conclusions and relevant references.

2. Hittites and Hittite Cuneiform Script:

The Hittites had used cuneiform signs to write about various topics. Van den Hout (2011) gives a classification of all available texts into genre. Some of them are historiography, treaties, edicts, instructions, loyalty oaths, laws, hymns and prayers, ritual scenarios, hippological texts and mythology written on wet clay tablets were baked and then later archived. Relatively very few of those tablets have been discovered and translated; most of them are still in the ground buried. Hittite cuneiform tablets that were from Corum Bogazkoy in Anatolia are in the memory of the world register by UNESCO in January 22, 2002.

In human history about 5000 years ago Sumerians discovered pictograph in Mesopotamia and many years later it evolved into another type of script which is called cuneiform that is brought to Anatolia by Akkadians and Assyrians during trading. By the time, Hittites used this cuneiform script and had later developed a script of their own called "Hittite cuneiform". In Hittite, clerks wrote cuneiform, using basic signs that form the cuneiform script, on wet clay tablets using sharp edged cane or reed or similar tools as stylus. After clerks wrote scripts on tablets, they baked tablets to become permanent and durable before archiving them. The Hittites were one of the first communities in the world history that had adopted the concept of archive-library.

Hittite cuneiform script has 375 different signs as noted by Ruster and Neu (1989). Gursel (1988) and Aktas and Gursel (1988) had shown that all these signs include five basic parts. Five basic signs in Hittite cuneiform script on tablets are given in Fig 2.1.

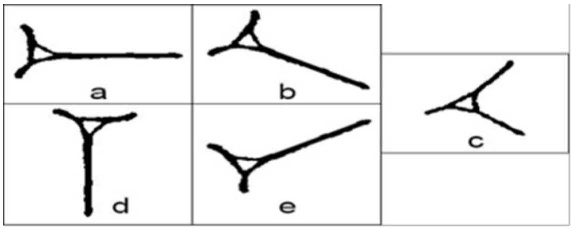


Fig 2.1 Basic signs in Hittite cuneiform script

Nearly more than 30 years ago, the first author supervised the first MS Thesis to recognize Hittite cuneiform signs using PROLOG programming language of that time (Gursel, 1988) and (Aktas and Gursel, 1988). That study had noted for the first time that Hittite cuneiform script consists of five basic signs given in Fig 2.1. Thus, the first study on computerized Hittite cuneiform signs in Turkey appeared at METU in 1988 (Gursel, 1988).

Such signs can represent a word or a syllable; also, a couple of them merge to represent a word. One of the basic signs is the horizontal sign (Fig 2.1a). Other basic signs were created by applying different angles (- 45°, - 90°, + 45°) to horizontal sign (Fig 2.1 b, d, e). Basic signs include also a different sign, which is named wedge (winkelhaken) (Fig 2.1c) written by pressing writing tool vertically to wet clay tablet.

In 1989 C. Ruster and E. Neu published a Hittite cuneiform sign dictionary named HZL (Hethitisches Zeichenlexikon) which includes Hittite cuneiform signs and their meanings. In HZL dictionary, sign number is index number of signs. This number is HZL number. Thus, in Hittite studies, HZL numbers refer to individual signs.

As noted earlier, B. Hrozny deciphered, for the first time in history, the following piece of text given as Fig 2.2 (Karasu, 2013).

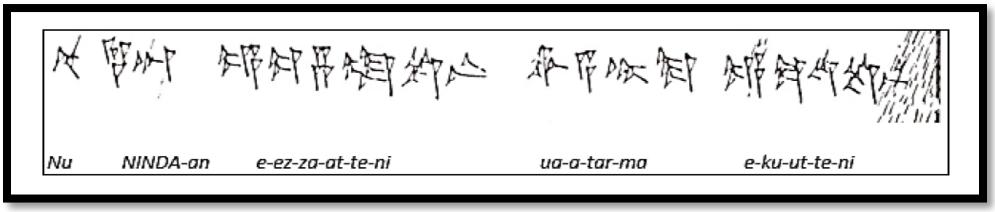


Fig 2.2 the first Hittite sentence translated into English: “You eat bread and drink water”

3. Related Other Works:

Referring to Van den Hout (2011) one notes, “At present there is no overview of Hittite literature written in English”. Apparently, there is one in German written by Haar (2006). Van den Hout also claims, “A systematic and up-to-date work on specifically Hittite art and archeology does not exist” Van den Hout (2011). Using ICT (Information and Communications Technologies) and especially the rapidly developing computer engineering tools and methodologies one might read and even perform translation on the Hittite cuneiform scripts hand copied already and even on unearthed original clay tablets. Another MS student, this time at Baskent University after more than twenty years, attempted to read cuneiform signs using an image processing technique (Dik, 2014). That study motivated another MS study for computerization of Hittite cuneiform text reading and translating using available fourteen different sign recognition algorithms and comparing their accuracies on various signs. In that study, also a brief data-mining application is made for combining scripts on fractured tablets using clustering algorithm of data-mining (Asuroglu, 2015). Another recent MS study at Baskent University Department of Computer Engineering devoted for a Knowledge-Based System or Expert System application on the previously digitally read cuneiform signs to extract their meaning in Hittite and later Turkish, German and English languages (Yesiltepe, 2015). Hittite cuneiform script is a collection of signs, therefore character recognition studies based on Chinese, Arabic, Japanese, Bangla and Tamil alphabet, in addition to Sumerian, Acadian and Assyrian cuneiform scripts, may be named as related work.

Dik (2015) made a study on the automatic translation of Hittite cuneiform signs. In this study, she developed a digital dictionary database, which included Hittite cuneiform signs and used an approach for Hittite cuneiform sign recognition by using Hausdorff Distance algorithm. She worked on the first Hittite sentence that Hrozny had solved (Karasu, 2013).

Tyndall (2012) applied data-mining algorithms to assemble transcripted cuneiform tablet parts that belong to a single tablet. He assigned the inventory number of tablet (given by Hittite experts) as class information, then broken parts matched by Hittite experts are assumed as single class, and dataset is created from these broken parts. During experiments, he used Naïve Bayes and Maximum Entropy classifiers and he computed classification performances.

Edan (2013) applied data-mining algorithms to Sumerian cuneiform signs. He acquired signs by a digital scanner and applied a pre-processing to reduce noise. Then, he created feature vectors, which consisted of horizontal and vertical distributions of cuneiform signs and number of connected components. He applied K-means clustering algorithm to find classes of cuneiform signs. After clustering, he applied Artificial Neural Network algorithm to cuneiform signs and classification performance was evaluated.

Rahma et al. (2006) proposed an algorithm called Intensity Curve to perform recognition of Sumerian cuneiform signs. In that algorithm first all signs were divided into equal horizontal partitions and in every partition pixel values and locations were calculated. After calculations, they transformed those values into a curve and local minimum values of curve created a feature vector. They applied the same procedures to vertical partitions too. They checked noisy, enlarged and reduced size versions of signs using a query database that holds original signs. They reported matching performance of Intensity Curve algorithm.

Ahmed (2012) proposed an algorithm called Symbol Structure Vector to perform recognition of Sumerian cuneiform signs. This algorithm starts with skeleton extraction of cuneiform signs. Features such as bending points and connection points of sign are also calculated, after skeleton extraction. A database contains features for later use. Real-time drawings of cuneiform signs are compared to sign database and matching performances were reported in the paper.

Sundar and John (2013) made a study on Tamil sign recognition. For every Tamil character, two different feature vectors were calculated. First of these vectors was calculated by using HOG algorithm, second one consists of geometric aspects of the sign. Using artificial neural network, they used these two feature vectors to compare and report the results as classification performance.

4. Hittite Cuneiform Sign Recognition:

4.1 Acquiring Digital Image of Hittite Cuneiform Signs

Portal Mainz website is used as the main source to acquire digital images of Hittite cuneiform signs. Portal Mainz is a website that is part of the Wurzburg University website. As shown in Fig 4.1, there are many Hittite cuneiform script tablet pictures available in the following website http://www.hethport.uni-wuerzburg.de/HPM/index.html).

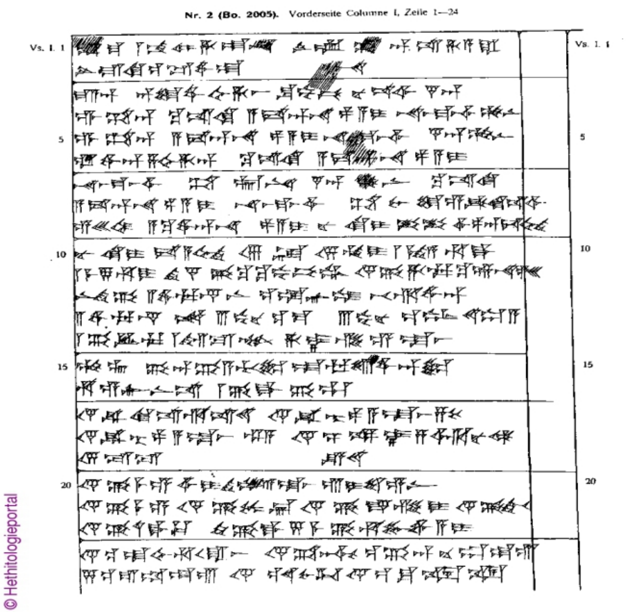


Fig 4.1 A copy of Hittite cuneiform script tablet

During the studies, summarized in this paper, the authors used these tablet pictures as a source for cuneiform signs.

In Portal Mainz website there is also a digital list created by Sylvie Vanseveren that includes all of Hittite cuneiform signs and their HZL index numbers. This list is referred to as 'V.S. digital sign' in those recent M.S. studies summarized in this paper. V.S. sign list includes high-resolution pictures of all Hittite cuneiform signs. Therefore, this list acts as a database for cuneiform signs in the M.S. studies. When finding the equivalent of signs in a tablet, V.S. digital list is used as a baseline for cuneiform signs.

4.2 Image Processing Algorithms for Hittite Cuneiform Sign Recognition

In the study by Asuroglu (2015) thirteen algorithms were used for computer based Hittite cuneiform sign recognition:

- B.U. Algorithm (Baskent University): Division of sign images into regions and calculation of an error rate (difference of number of black pixels in every region).
- MATLAB Regionprops library. This library helps to calculate geometric features of an image. http://www.mathworks.com/help/images/ref/regionprops.html
- SIFT algorithm (Scale Invariant Feature Transform) (Lowe 2004).
- SURF algorithm (Speeded up Robust Features) (Herbert et al., 2006).
- FAST algorithm for Corner Detection (Features from Accelerated Segment Test) (Rosten and Drummond, 2006).
- BRISK algorithm (Binary Robust Invariant Scalable Keypoints)

(Leutenegger et al., 2011).
- MSER algorithm (Maximally Stable Extremal Regions) (Matas et al., 2002).
- ORB algorithm (Oriented FAST and Rotated BRIEF) (Rublee et al., 2011).
- HARRIS corner detection algorithm (Harris and Stephens, 1988).
- Hausdorff Distance algorithm: When comparing two signs, distances between these two signs are calculated and minimum distance is selected (Huttenlocher et al., 1993).
- Calculation of structural features using Hough transform (Chunhavittayatera et al., 2006).
- Hierarchial Centroid (H.C.) algorithm: Division of sign image into partitions and centroids of every partition are extracted as a feature. (Armon, 2011 and Faiganbaum et al. 2016).
- HOG (Histogram of Oriented Gradients) algorithm (Dalal and Triggs, 2005).

Some of these algorithms were derived using functions that belong to the MATLAB Toolbox (e.g. Algorithm 1). Another example is algorithm 2 that belongs to the MATLAB Library. Algorithms like 3, 4 and 5 belong to OpenCV Library (http://opencv.org/)

5. Data Mining Examples on Hittite Cuneiform Signs:

In Hittite cuneiform script, there are many geometrically similar signs. Thinking of gathering these similar signs in different categories has created data-mining view of this study. During the study, geometric features were defined first and categorization of geometrically similar signs was carried out by K-means clustering algorithm, which is a popular data-mining algorithm (Han and Kamber, 2006; Ahamed and Hareesha, 2012). After categorization, popular data-mining classification algorithms are applied on the Hittite cuneiform signs and classification performances are reported in the following subsections.

5.1 Hittite Cuneiform Signs Dataset

In data- mining examples, dataset consists of geometric features of Hittite cuneiform signs. These cuneiform signs were selected from V.S. digital list. Digital image acquisition phase of cuneiform signs is the same as Subsection 4.1 of the paper. Geometric features are extracted by Algorithm 2 of MATLAB Regionprops library. These geometric features are Area, X coordinate of centroid, Y coordinate of centroid, Euler Number, Extent, Eccentricity and EquivDiameter. Geometric features are extracted for every cuneiform signs that are used in data-mining algorithms. Finally, a dataset with 7 features is constructed.

5.2 Data-mining Algorithms That Were Used in Hittite Cuneiform Signs

5.2.1 K-means clustering algorithm

K-means clustering algorithm is an algorithm of data-mining that has descriptive model structure. It is used for assigning class labels to data that class labels are unknown. K-means is one of the most popular data-mining clustering algorithms because it can be easily implemented and does not take too much processor time (Armon, 2011). Main purpose of K-means is to divide unlabeled data to K class by using features of data. Algorithm places data to a feature space and make clustering on this feature space.

5.2.2 J48 decision tree classification algorithm

J48 decision tree algorithm is the www.cs.waikato.ac.nz/ml/weka implementation of Quinlan's C 4.5 (Quinlan, 1993) decision tree algorithm (Sharma and Sahni, 2011).

5.2.3 k-Nearest Neighbor (kNN) classification algorithm

k-nearest neighbor (kNN) algorithm was proposed by Cover and Hart (1967). Algorithm is used in many areas; reasons behind such popularity are fast classification model building and good classification results on noisy data (Bhatia, 2010). Algorithm works with principle of "Classify according to nearest neighbors" (Patel and Patel, 2016).

5.2.4 Artificial Neural Network (ANN) classification algorithm

Artificial neural network is applicable in many areas including finance, engineering, geology and physics (Suguna and Thanushkodi, 2010) and (Pradhan and, S. Lee, 2007). ANN structure models human brain's most important aspects, which are learning, interpretation of information and inference. ANN developed to perform these processes automatically. ANN's mathematical model of decision and learning process are inspired by human brain.

6. Development of a Sample Rule Tree for Hittite Language:

6.1 General

As mentioned earlier, the Hittite language belongs to Indo-European language family. That family covers a large geographic area in the world. For this reason, certain differences have grown among themselves in the languages in the same family. Hittite language is the oldest Anatolian branch of the family. It therefore attracts attention of linguists in various countries (Alkan, 2011 and Arikan, 1998).

6.2 Some Basic Properties of Hittite Language

Hittite language is based on syllable structure similar to other old Middle Eastern languages. Hittite language has 375 signs that may be syllables, ideogram and numbers.

In Hittite cueiform script signs there are some Accadian and Sumerian words too. Figure 6.1 shows a tree diagram to differentiate these words.

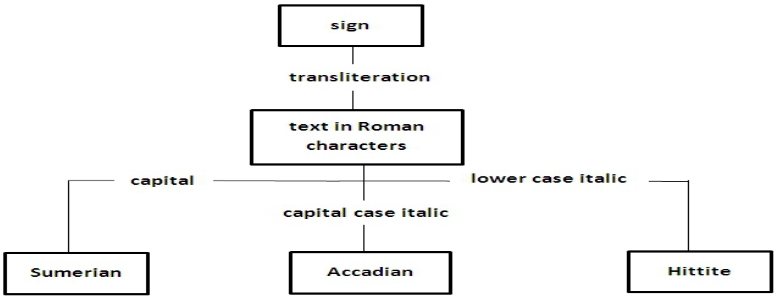


Fig. 6.1 Basic tree structure for a cuneiform sign

6.3 Expert System Rules Base on Hittite Grammar Rules

It is not possible to summarize the very rich grammar rules of a language like Hittite. Therefore the tree structure given as Fig. 6.2 will serve as an introduction to the rule formulation of Hittite grammar to apply on an expert system shell.

Hittite language has 3 basic syllables as shown in Table 6.1

Table 6.1 Hittite Syllable Types

Kind of Syllable	Type of Syllable	Example
Simple Syllable	Consonant+ vowel	bu pu
Simple Syllable	Vowel+consonant	ar er
Complex Syllable	Consonant+vowel+consonant	gal lal

Further details of Hittite grammer is given in references (Karasu,2013), (Van den Hout,2011), (Arıkan, 1998) and (Hoffner and Melchert,2013).

7. A Hittite Information System Proposal

Translation of Hittite Texts on clay tablets written in cuneiform scripts is a tedious and highly expertise needed work. After digital image processing of signs, the needed work may be summarized in three basic steps as follows:

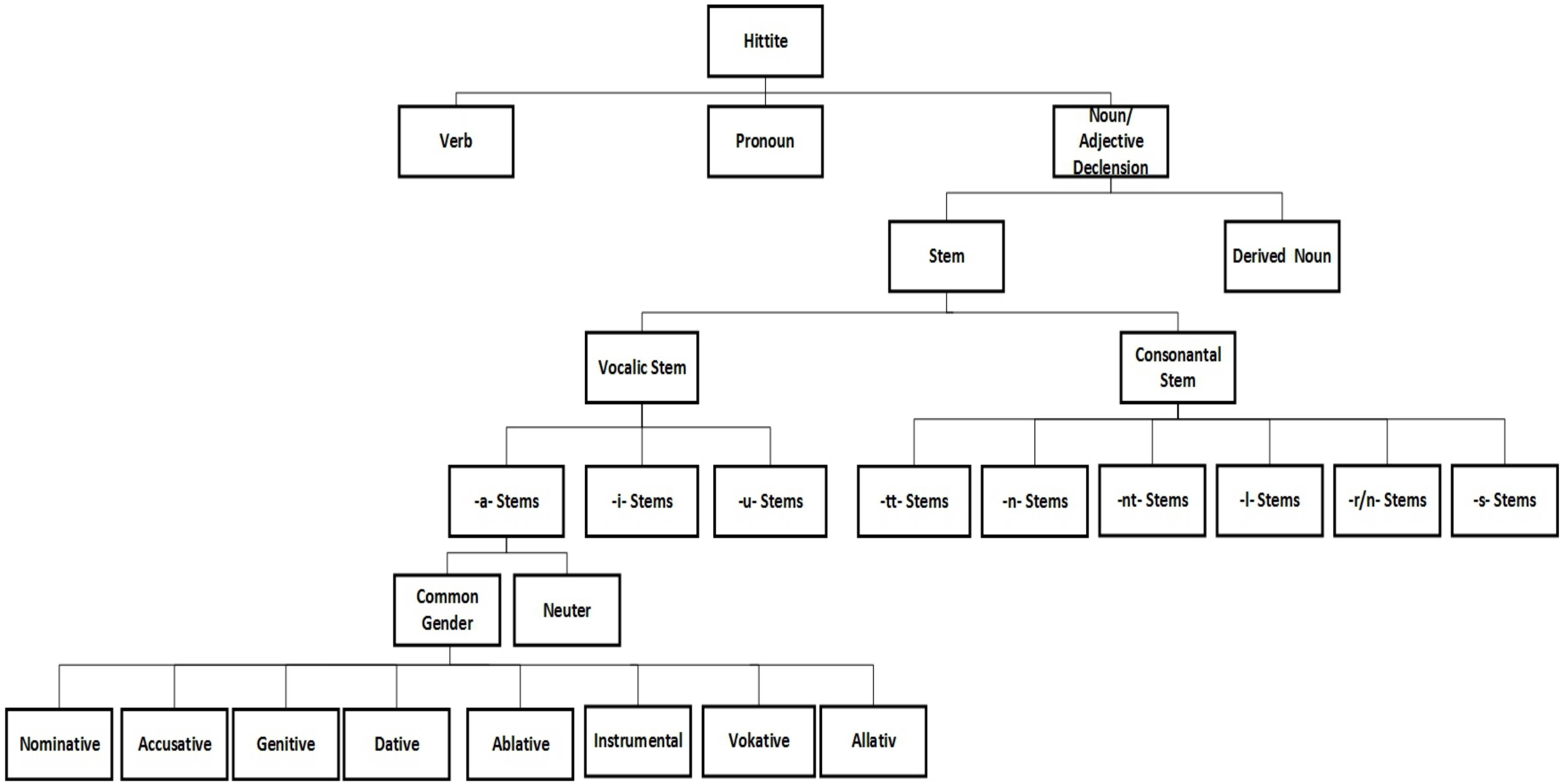


Figure 6.2 Hittite Grammar rule tree for a noun

a) Transliteration
b) Transcription
c) Translation

Referring to (Van den Hout, 2011) p.11, "we call the process of transferring a cuneiform text to Latin alphabet as transliteration noting the differences between Hittite cuneiform signs, Sumerograms and Akadogram."

Everything Hittite is in lower case, each individual cuneiform sign separated by hyphens (e.g. is-ha-as), Sumerograms are given in roman capitals (e.g. EN), and a series of Sumerian word signs is separated by periods (e.g. MUNUS. LUGAL) meaning "woman, king" that is "queen". Akadograms are also capitalized but italicized and hyphenated (e.g. *U-UL* i.e. "not" or *BE-LU* i.e. "lord").

The next process after transliteration is transcription which means an attempt at making real words out of the transliterated sign sequences. In transcription the symbol = is often used. It indicates the so called "morpheme boundaries." Morphemes are the smallest meaningful grammatical elements (Van den Hout, 2011) p.13.

The Hittite language has four vowels: /a, e, i, u/. There is no /o/. Order of Hittite alphabet is given as follows: a e h i k/g l m n p/b s t/d u w z

The last basic step was stated as translation. Transcripted text is translated into living languages of Turkish, German, English and others. Especially, for this process a human expert or an expert system or a knowledge-based system having the grammar rules of Hittite language are needed badly.

In a proposed information system, starting with a computer based reading of Hittite cuneiform signs on clay tablets and going through all the steps until finishing and publishing translation is depicted in Fig. 7.1. The basic processes of the proposed information system is modeled using DFD (Data Flow Diagram) technique (e.g. Aktas, 1987; Braude and Bernstein, 2011; and Schach, 2011).

The DFD - Overview Diagram given as Fig. 7.1, has the following seven processes:

Process 1.Get digital image of the clay tablet in museum or archeological site;

Process 2. Process digital image using sign recognition algorithms to get digital image of the script;

Process 3. Transliterate text image using computer;

Process 4. Transcript text image using computer;

Process 5. Translate text using computer;

Process 6. Let human expert(s) refine the computer-based translation of the text;

Process 7. Share translation in academia.

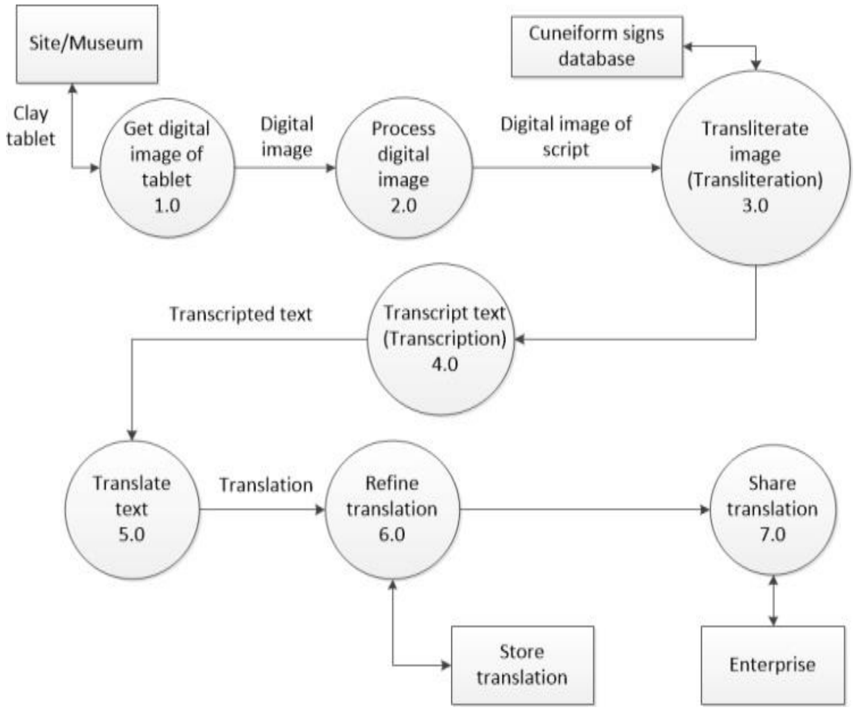


Fig 7.1 DFD of the proposed Hittite information system

8. Expert Systems

8.1 General

Expert Systems or Knowledge-Based Systems is a product of Artificial Intelligence (AI) that started in 1950s. AI is using the computers to exhibit human-like cognitive and decision-making capabilities that are human intelligence.

Referring to (Becerra-Fernandez, 2004) and (Awad and Ghaziri, 2004) to define Knowledge-Based Systems or Expert Systems, one may state that a Knowledge Based System preserves and apply human expertise on any particular area. It is also known as "Knowledge Engineering" (Becerra-Fernandez, 2004; Aktas and Cetin, 2011). A Knowledge-Based System Developers (KBS) may be defined according to point of views of End Users and Developers. From end users perspective, a KBS has three components such as the intelligent program, the user interface, and a problem-specific database as depicted in Fig. 8.1. The Intelligent program is the main part of KBS from the stand point of a user. It solves the users' problems. It is like a black box to user. Using the User Interface, user can control the system in solving his/her problem(s). The last component, namely Workspace, is a problem specific database where the system reads any inputs and writes its outputs.

Knowledge Engineers (KEs) are the developers of a KBS. From a KE's view, a KBS has two major components as the Intelligent Program and the KBS Development Shell as shown in Fig. 8.2.

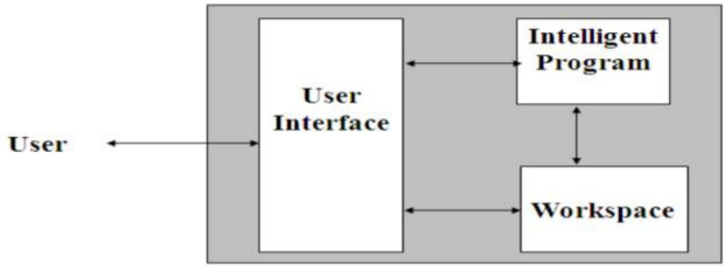


Fig. 8.1 End users point of view for KBS

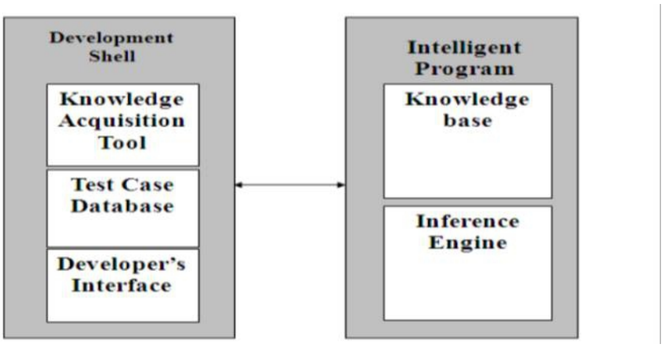


Fig 8.2 Developers point of view for KBS

Intelligent Program is the same as a user sees it as in Fig. 8.1. KE now sees its components as Knowledge Base and Inference Engine. Knowledge Base contains the knowledge used by the system and Inference Engine provides the functionality to implement the automated reasoning in solving the problem. KBS shell or development environment also called in Fig. 8.2 has set of tools for creation of knowledge in the Intelligent Program, such as Knowledge Acquisition Tool, Developers' Interface and a Test Case Database. The Knowledge Acquisition Tool assists the KE in the construction of the Knowledge Base component of the Intelligent Program. During development, KE interacts with the human experts of the problem domain and acquire knowledge from them to keep in the Knowledge Base.

The second component, the Test Case Database contains sample problems executed successfully earlier in the KBS. Whenever a change in the knowledge-base is made one can execute these test cases to verify that these benchmark test cases are still solved correctly.

9. A Proposed Expert System

It appeared to the authors that using an available expert system shell would provide a good support in getting the meaning of Hittite cuneiform signs. As noted earlier, grammar rules of Hittite language are already available (Karasu, 2013; Van den Hout, 2011; Hoffner and Melchert, 2013; Unal, 2007). Using the grammar rules available one could develop IF... THEN rules to be stored in the Knowledge Base of an expert system. In order to demonstrate the possibility of that idea a commercially available expert system named Exsys CORVID is used successfully. Hittite language has a very rich grammar rules. It has a highly conservative verbal system and rich nominal declension. As noted earlier few times, the language is written in cuneiform and it is one of the earliest examples of Indo-European language family other than Vedic Sanskrit.

It is impractical to include all these rules in a tree structure. In Fig 4.2 a rough tree structure of Hittite grammar was already given. The Hittite nominal system has the following cases: nominative, accusative, dative-locative, genitive, allative, ablative, and instrumental, and distinguishes between two numbers (singular and plural) and two genders, common (animate) and neuter (inanimate). The distinction between genders is rudimentary, with a distinction generally made only in the nominative case, and the same noun may be used for both genders. Considering a Hittite noun say, "antuḫša", which means man, human being or person in English language, its declension is given as Table 9.1 (Karasu, 2013). Fig 9.1 is prepared to summarize the grammar rules of the noun/adjective declension in Hittite language to show how a noun may be placed on a tree so that IF... THEN rules can be generated to be placed in the expert system shell of the knowledge base. The grammar rules given in Table 9.1 for a Hittite word were transformed into IF... THEN rules given as Figure 9.1 to be stored into EXSYS Corvid software.

Table 9.1 A noun "antuḫša" "man" declension example

	Cases	**Singular (sg.)**
antuḫša "man"	Nominative Common	antuḫšaš
	Accusative	antuḫšan
	Nominative-Accusative n.	-
	Vocative	-
	Genitive	antuḫšaš
	Dative- Locative	antuḫši
	Allative	-
	Ablative	antuḫšaz
	Instrumental	antuḫšet

IF antuḫšaš is man
THEN man_sg_nominative_com
IF antuḫšan is man
THEN man_sg_accusative
IF antuḫšaš is man
THEN man_sg_genitive
IF antuḫši is man
THEN man_sg_dative-locative
IF antuḫšaz is man
THEN man_sg_ablative
IF antuḫšet is man
THEN man_sg_instrumental

Fig 9.1 Logical IF… THEN rules for the sample noun

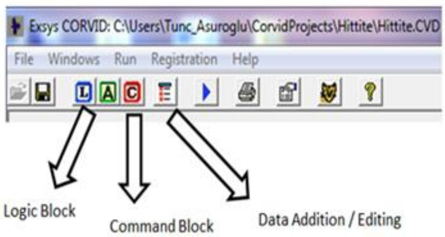
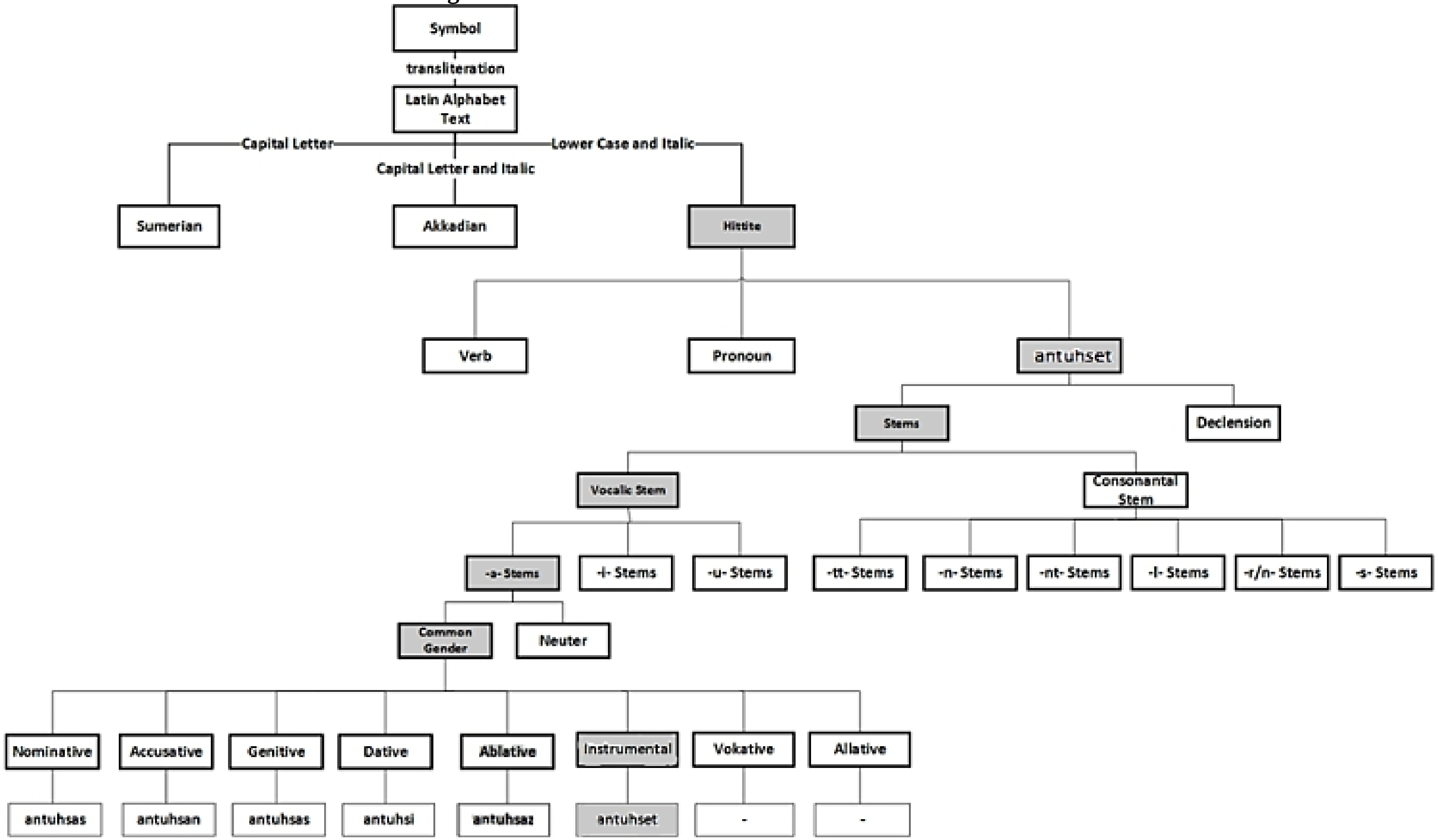


Fig 9.3 Overview of an ExsysCORVID Application

Logical IF… THEN rules are placed in Fig 6.2 to get new figure as Fig 9.2. Fig 9.3 is an overview of ExsysCORVID Application.

Figure 9.2 Traverse of the Hittite Grammar tree for a noun



The placement of the rules in the expert system is given as Fig. 9.4.

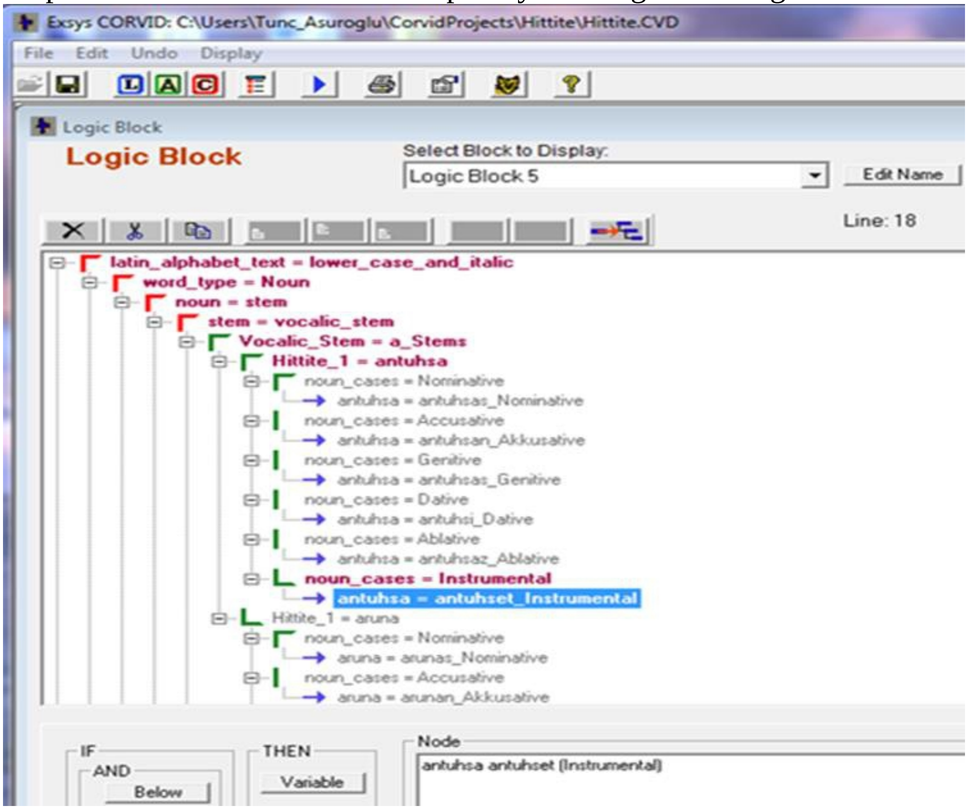


Fig 9.4 The placement of the rules in the Expert System

Rules are summarized in the Rule View of Exsys CORVID in Fig. 9.5. The result is also shown there.

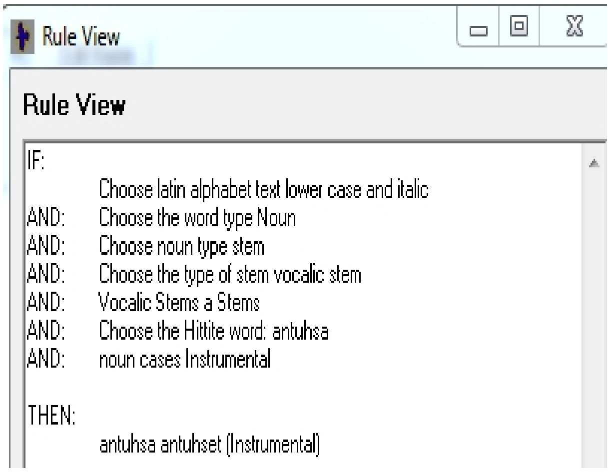


Fig 9.5 Rule View of Exsys

10. Summary, Conclusions and Extensions of the Study

10.1 Summary

In this paper a proposed computer-based information system project is summarized. The major objective of that project is to read Hittite Texts in cuneiform scripts on clay tablets or their hand-copied versions or photographs using computer based sign recognition techniques. Developing IF . . . THEN logical rules of Hittite language grammar to be loaded into an expert system (or knowledge-based system) would eventually be used in translating Hittite Texts into first Turkish and later into English and German languages.

In three consecutive M.S. studies completed during the last few years at Computer Engineering Department of Başkent University-Ankara, Turkey, thirteen computer based sign recognition algorithms have been successfully used to read Hittite cuneiform signs comparing their accuracy and speed. Next, studying some of the Hittite language grammar rules, they were converted into IF . . . THEN rules to be loaded into an existing expert system shell. That system was then used to convert Hittite cuneiform signs into Turkish words. ANN algorithm was also used in a study to deal with fractured clay tablets.

It is clear that such a project as a whole is a very complicated and complex task. It requires more time, expertise and money. One may then consider this paper as a Summary of technical feasibility study of the whole project in which sign recognition is successfully performed and grammar rules of Hittite language are converted into logical rules to be stored into expert system shells for translation into Turkish and later into English and German languages. These are all successfully accomplished during these recent studies. Thus the first phase of the whole project is finished to prove that it can be done. In a way one may claim that a prototype of the whole project is successfully finished.

10.2 Conclusions

In realization of this phase of the project, Open Access sign recognition algorithms together with MATLAB software were used. In order to demonstrate the applicability of an expert system, trial version of a Professional expert system shell named ExsysCORVID was used successfully.

For the final phase of the future project an object oriented software development methodology named RUP (Rational Unified Process) or more recent OpenUP would be used. For the time being existing application may be termed as Open Access Approach.

There are some dictionaries available to translate Hittite signs into Turkish and English (Güterbock, Hoffner &Van den Hout-Chicago Project) or German (Ruster & Neu, 1989). The Chicago Project is known as OCHRE (Online Cultural Heritage Research Environment) and during recent years its web-based copy has been developed under the name eCHD (http://ochre.lib.uchicago.edu/eCHD/).

In this early phase of the project, the expected benefits of the project were defined as Fast Response, High Performance, Understandability and Reliability. Considering the final project, the additional expected benefits may be cited as: Safety, Availability, Cost, Maintainability, Time, Energy consumption, Usability and Productivity as suggested by Gomathy & Rajalakshmi (2014).

10.3 Extensions of the Study

As stated earlier, the major Objective of this paper is to prove the technical feasibility of a large project; in a way to serve as a prototype project summary. The whole project will take definitely more time, more human resources and more money.

As noted in the text of the paper, ANN (Artificial Neural Network) algorithm was also used in one of the theses. As an extension, Deep Learning techniques and CNN algorithm may also be used to repair cuneiform scripts of fractured tablets.

Subecz (2019) just very recently noted that the detection and analysis of events in natural language applications play a significant role. That point would be taken into consideration in the final project.

Another interesting approach to be tried in future may be the use of 2D&3D document formats for cuneiform script applications (Bogacz, Massa & Mara, 2015) and (Rothacker. Fisseler, Müller, Weichert &Fink, 2015).

The output is planned to be in Turkish first. At the end of the whole project, English and German translations would also be included.

Acknowledgment

The authors would like to express their gratitude to Dr. C.K. Gomothy and an anonymous Reviewer for their very careful reviews and supportive comments. Thanks are extended to Prof. Dr. I. C. Karasu and Prof. Dr. Y. Arikan for their kind support and encouragement during the project “Hittite Cuneiform Script and Further: Hittite Information System project (HIT-IS Project)”. Thanks are also extended to Asst. Prof. Dr. Tulin E. Ayyildiz, Asst. Prof. Dr. E. Sumer, Asst. Prof. Dr. M. Sert, Miss M.B. Yesiltepe, Miss S. Sahin, Mr. S. Isik and Mrs. N. Ugur for their contributions in the project. We also owe thanks to Mrs. D. Ölçer and Asst.Prof.Dr. D. Şener for their support. The corresponding author for the paper is Prof. Dr. A. Z. Aktas. University copy of MATLAB software, Visio Software of Microsoft and Trial version of Exsys CORVİD were used in the study; we thank cordially to the Companies.

References:

1. Karasu,C.(2013).Hititçe ve Hitit çivi yazısı. Editörler: M. D. Alparslan ve M. Alparslan, Hititler, Yapı Kredi Yayınları, ss.84 – 93 (in Turkish).
2. Dik, E.C.(2014) Hitit çiviyazısı işaretlerinin otomatik çevirisi. Yüksek Lisans Tezi, Başkent Üniversitesi, supervised by Prof. Dr. A. Z. Aktas (in Turkish).
3. Asuroglu, T.(2015). Hitit çiviyazisi işaretlerinin bilgisayar desteği ile okunmasi ve veri madenciliği uygulama örnekleri. Yüksek Lisans Tezi, Başkent Üniversitesi, supervised by Prof. Dr. A. Z. Aktas (in Turkish).
4. Yesiltepe, B.M.(2015). Hitit çiviyazılı metinlerin okunmasında uzman sistem uygulama örnekleri. Yüksek Lisans Tezi, Başkent Üniversitesi, supervised by Prof. Dr. A. Z. Aktas (in Turkish).
5. Van den Hout,T.(2011). The elements of Hittite.Cambridge University Press.
6. Ruster, C. &Neu, E. (1989). Hethitisches zeichen lexikon: inventar und interpretation der keil schriftzeichen aus den Bogazkoy-Texten. O. Harrassowitz.
7. Gursel,H. (1988). An expert system for cuneiform interpretation. M.S. thesis, Comp. Eng. Dept. METU, Ankara.
8. Aktas,Z.&Gursel,H.(1988).Çiviyazısı metinlerin uzman sistemler yardımıyla çözülmesi.5.TürkiyeBilgisayar Kongresi,Ankara,(in Turkish).
9. Haar,V.(2006). DieHethitische literatur. De Gruyter, April 1.
10. Tyndall, S.(2012). Toward automatically assembling Hittite-language cuneiform tablet fragments into larger texts. 50th Annual Meeting of the Association for Computational Linguistics,vol.2, pp. 243-247.
11. Edan, N.M. (2013). Cuneiform symbols recognition based on K-means and neural network. Journal of Computational Mathematics, vol.10, no.1, pp. 195-202.
12. Rahma, A. M., Yousif, H.& H. Alani(2006). Cuneiform symbols recognition using intensity curves. The International Arab Journal of Information Technology, vol.3, no.3, pp. 237-241.
13. Ahmed, K. K.(2012).Online Sumerians cuneiform detection based on symbol structural vector algorithm. Journal of the College of Education for Women, vol.23, no.2, pp. 545-553.
14. Sundar, K.A. ,&John M.A.(2013). High precision printed character recognition method for Tamil script. IEEE International Conference on Computational Intelligence and Computing Research (ICCIC).
15. http://www.hethportuni-wuerzburg.de /cuneifont/
16. http://www.mathworks.com/help

17. Lowe, D. G. (2004). Distinctive image features from scale-invariant key points. International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110.
18. Herbert, B., Tuytelaars, T. &Van Gool, L. (2006) SURF: speeded up robust features. Lecture Notes in Computer Science Volume 3951, pp. 404-417.
19. Rosten, J. E. & Drummond,T.(2006) Machine learning for high-speed corner detection. 8th European conference on Computer Vision (ECCV), pp. 430-443.
20. Leutenegger, S., Chli, M.&Siegwart, R.Y.(2011). BRISK: Binary robust invariant scalable keypoints. IEEE International Conference on Computer Vision , pp.2548-2555.
21. Matas, J., Chum, O., Urban, M. & Pajdla, T.(2002) Robust wide-baseline stereo from maximally stable extremal regions. British Machine Vision Conference, pp. 384-393.
22. Rublee,R., Rabaud,V., Konolige,K.& Bradski,G.(2011).ORB: an efficient alternative to SIFT or SURF. IEEE International Conference in Computer Vision (ICCV), pp. 2564-2571.
23. Harris,C.&Stephens, M.(1988). A combined corner and edge detector. Alvey Vision Conference, pp. 147-151.
24. Huttenlocher D. P,. , Klanderman, G. A. & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.15, no.9, pp. 850-863.
25. Chunhavittayatera, S., Chitsobhuk, O. & Tongprasert, K.(2006). Image Registration using Hough Transform and Phase Correlation", International Conference on Advanced Communication Technology (ICACT), pp. 973-977.
26. Chitsobhuk, S.O. & Tongprasert, K.(2006). Image registration using Hough transform and phase correlation. International Conference on Advanced Communication Technology (ICACT), pp. 973-977.
27. Armon, S. (2011). Handwriting recognition and fast retrieval for Hebrew historical manuscripts. M.S. thesis, Hebrew University, Jerusalem, Israel.
28. Faigenbaum-Golovin, S., Shaus, Sober, A., Finkelstein, I., Levin, D., Na'aman, N., Sass, B., Turkel, E. & Piasetzky, E. (2016). Algorithmic handwriting analysis of Judah's military correspondence shed light on composition of biblical texts. Proc. of the National Academy of Sciences of Israel, 113(17) DOI: 10.1073.
29. Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. IEEE Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886-893.

30. http://opencv.org/
31. Han, J. & Kamber, M. (2006). Data-mining concept and techniques. 2nd Edition, Elsevier Inc.
32. Ahamed, S. &Hareesha, K. (2012). Dynamic clustering of data with modified k-means algorithm. International Conference on Information and Computer Networks (ICICN), pp. 221-225.
33. www.cs.waikato.ac.nz/ml/weka/
34. Quinlan, R. (1993). C4.5: Programs for machine learning. Morgan Kaufmann Publishers.
35. K. Sharma and S. Sahni, (2011) A comparative study of classification algorithms for spam email data analysis. International Journal on Computer Science and Engineering (IJCSE), vol.3, no.5, pp.1890-1895.
36. Cover, T. & Hart, J. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, vol.3, no.1, pp.21- 27.
37. Bhatia, N. (2010) Survey of Nearest Neighbor Techniques. International Journal of Computer Science and Information Security, vol.8, no.2, pp. 302-305.
38. Suguna, N. & Thanushkodi, K. (2010) An improved k-nearest neighbor classification using genetic algorithm. International Journal of Computer Science Issues, vol.7, Issue 4, no.2, pp. 18-21.
39. Pradhan, B. & Lee, S. (2007). Utilization of optical remote sensing data and GIS tools for regional landslide hazard analysis using an artificial neural network model. Earth Science Frontiers, vol.14, no.6, pp. 143-152.
40. Arikan, Y. (1998). Hitit dini üzerine bir inceleme. Ankara Üniversitesi Dil ve Tarih- Coğrafya Fakültesi Dergisi Cilt 38, Sayı 1-2. (In Turkish).
41. Hoffner, H. A. & Melchert, H. C. (2013). A Grammar of the Hittite Language. Student Loose Edition.
42. Aktas, A. Z. (1987). Structured analysis and design of information systems. Chapters 1-4. Prentice Hall, 1987.
43. Braude, E.J. & Bernstein, M. E. (2011) Software engineering. Wiley, 2nd Ed.
44. Schach, S. R. (2011). Object-Oriented and Classical software engineering. McGraw Hill, 8th Ed.
45. Becerra-Fernandez, I., Gonzalez, A. & Sabherwal, R. (2004). Knowledge Management. Pearson/ Prentice-Hall.
46. Awad, E. &. Ghaziri, H. M. (2004). Knowledge Management. Pearson / Prentice Hall.

47. Aktas, A.Z. & Cetin, S. (2011). Rebirth of a Discipline: “Knowledge Engineering. CMES (Computer Modeling in Engineering and Science), Vol. 1943, no.1, pp.1-29.
48. Ünal, A. (2007). A Concise Multilingual Hittite Dictionary, Verlag Dr.Kovacˇ, Hamburg.
49. Güterbock, Hoffner &Van den Hout (existing).The Hittite Dictionary. The Oriental Institute of the University of Chicago, USA.
50. Gomathy, C.K. & Rajalakshmi, S. (2014). European Scientific Journal, March 2014 edition vol.10, No.9 ISSN: 1857-7881 (Print) e-ISSN 1857-7431 183.
51. Subecz, Z.(2019).Event Detection and Classification in Hungarian Natural Texts. European Scientific Journal July 2019 edition Vol.15, No.21 ISSN: 1857 – 7881 (Print) e - ISSN 1857- 7431 411.
52. Bogacz, B., Massa, J. & Mara, H. (2015). Homogenization of 2D&3D Document Formats for Cuneiform Script Analysis.HIP’15 Proceedings of the 3rd International Workshop on Historical Document Imaging and Processing, pp.115-122, Gammarth, Tunisia, August22, 2015.
53. Rothacker, D., Fisseler, D., Müller, G.G.W., Weichert, F. & Fink, G.A. (2015). Retrieving Cuneiform Structures in a Segmentation-free Word Spotting Framework. Workshop on Historical Document Imaging and Processing.