From Ancient Egyptian Language to Future Conceptual Modeling

Peter P. Chen

Computer Science Department
Louisiana State University
Baton Rouge, LA 70803, USA
chen@bit.csc.lsu.edu

Abstract. This paper discusses the construction principles of ancient Egyptian hieroglyphs from the point of view of conceptual modeling. The paper starts with a summary of author's previous work on the correspondence between the Entity-Relationship diagrammatic (ERD) technique and two natural languages: English and Chinese. In one previous work, the similarity between the English sentence structure/grammar and the construction blocks and rules of ERD was discovered. In another work, the similarity between the Chinese character construction methods and the ER modeling principles was also discovered. In this paper, construction methods of the ancient Egyptian hieroglyph are analyzed with respect to the construction methods of Chinese characters and the conceptual modeling principles. At the end, possible applications and extensions of this research work are discussed.

1 Introduction

One of the most crucial steps in software engineering is to model the user requirements correctly. An accurate model can help not only in developing better quality software system but also in speeding up the software development process, which usually overran in costs and time. More and more systems analysis and design methodologies rely on graphical representations of the user requirements and the software systems. The most frequently used diagrammatic techniques are Data Flow diagrams, Entity-Relationship Diagrams (ERD), Object Diagrams, and State Transition Diagrams. Besides software engineering, ERD and semantic networks are used widely in data and knowledge engineering. Detailed descriptions of these techniques can be found in many software engineering and systems analysis books (for example, [20]) or data/knowledge engineering books.

The Entity-Relationship (ER) diagram technique [2] is a graphical tool for information system designers and users in describing and communicating their understanding of the world. Since its formal introduction in 1976, it has become one of the most widely used techniques in systems analysis and database design projects and the cornerstone of the CASE (computer-aided software engineering) industry [15, 17]. Although the ER diagram technique is primarily used by computer professionals

and users, its use has been spreading to other fields such as accounting and music composition. Many recent papers on ER model extensions and applications can be found in [1, 12, 13, 14, 18]. For more theoretical work, the reader can refer to [19, 21].

On the other hand, natural languages are the daily tools for the general public in describing and communicating their understanding of the world. Because both the ER diagram (ERD) techniques and the natural languages satisfy similar human needs, these two "human communication" techniques should have something in common. Furthermore, if we can understand better the correspondence between "ERD" and "natural language," it is very likely that we can use this knowledge to improve our modeling methodologies, techniques, and tools.

In this paper, we will first start with a review of key concepts between "English grammar constructs" and "ER diagram constructs." This section is a summary of the author's previous work on this topic. Then, the next section summarizes the author's previous work on the principles for constructing Chinese characters. Some of these principles are well known by those familiar with the Chinese written language, but we modify the concepts to fit better with conceptual modeling professionals. The rest of the principles either are not well known or are proposed by the author. The fourth section is a short introduction to the history of ancient Egyptian languages. The fifth section tries to organize the construction principles of ancient Egyptian language hieroglyphs into the same framework of modeling principles as the construction of Chinese characters. The final section states the conclusions and the future research/application directions. Throughout this paper, we assume that the reader has some basic understanding of the notations and concepts of the ERD technique.

2 Review of the Correspondence between English Sentence Structures and ERD Constructs

The correspondence between the English sentence structures and the ERD construct was first presented at the 2nd ER Conference in 1981 in Washington, D.C., and later published in [3]. A summary of the basic translation rules is summarized in Table 1. For example, a "common noun" (such as "desk," "car") in English is a possible candidate for an entity type in an ERD. A "proper noun" (such as "George Washington") is a possible candidate for an entity (an instance of an entity type) in an ERD.

English Grammar Structure	ERD Structure
Common Noun	Entity Type (a possible candidate)
Proper Noun	Entity (candidate)
Transitive verb	Relationship type (candidate)
Intransitive verb	Attribute type (candidate)
Adjective	Attribute for entity
Adverb	Attribute for relationship
Gerund (a noun converted from a verb)	An entity type converted from a relationship type
Clause	A high-level entity type which hides a detailed ERD

Table 1. Correspondence between English sentence structures and ERD constructs

This technique can be used in several ways:

- As an early stage requirement analysis tool: it can help users to identity entity types, relationship types, attributes, and high-level ERD constructs based on the English sentence structures. Recently, researchers in OO (Object-Oriented) Analysis methods also started to advocate the use of English "nouns" as a way to identify possible "objects," and this is the same position we advocated in [3].
- As the basis for (manually or semi-automatically) converting a large amount of requirements specification documents (in English) into ERD-like specifications. Several large consulting firms are practicing this technique. This technique can also be used in reverse direction, that is, to use ERD in assisting users to formulate a more English-like query. In other words, to use it as a basis for building a more English-like interface to database management systems.

In the following section, we will discussions the correspondence between Chinese characters and ERD constructs.

3 The Construction Principles of Chinese Characters

Chinese written language is one of the earliest written languages in the world. The earliest development of Chinese characters was claimed to start as early as 8,000 years ago when a few picture-like characters were carved onto turtleback shells. However, a relatively useful set of approximately 4,500 characters was developed much later (approximately around 2,300 years ago) in the late Shang dynasty.

Most Western language characters (and even some modern Asian language characters) are phonetic-based, while Chinese characters are mostly picture-based. Chinese characters are ideograms; each one represents an idea, a thing, etc. Today, there are tens of thousand Chinese characters in circulation. How can a human brain store and retrieve so many ideograms? Fortunately, there are organized methods for memorizing some basic characters and for constructing new characters from existing ones.

Even though the various methods for constructing Chinese characters have been known for a long time and the genealogy of many Chinese characters have been investigated by researchers (see, for example, [11]), virtually none had looked at the

construction methods from the viewpoint of conceptual modeling. In our study [8], we identified the following principles for Chinese character construction, which are also commonly used in conceptual modeling:

- Principle of Physical Resemblance
- Principle of Subset
- Principle of Grouping
- Principle of Composition (Aggregation)
- Principle of Commonality
- Principle of an-instance-of (Something).

We will explain each of these principles in Section 5 when we discuss the construction methods of ancient Egyptian hieroglyphs.

4 History of Ancient Egyptian Written Language

Similar to the Chinese, the Egyptians developed relatively sophisticated written language about four to five thousand years ago. Ancient Egyptian language can be divided into five different phases [9, 10, 16]:

1. Old Egyptian (approximately 2700-2200 BC)
2. Middle Egyptian (approximately 2200-1800 BC)
3. New Egyptian (approximately 1580-700 BC)
4. Demotic (approximately 700 BC-600 AD)
5. Coptic (600 AD - 1,000 AD).

The hieroglyphic script is of a pictorial written form of the Ancient Egyptian language. The earliest documented occurrence dates back to the pre-dynastic period, and the latest occurrence dates back to approximately 400 AD. At the end, the number of hieroglyphs grew to approximately 6,000, which was sufficient to express some complicated thoughts and events during that time.

5 Principles on Constructing an Ancient Egyptian Hieroglyphs

In the following, we will discuss the construction methods of ancient Egyptian hieroglyphs using the same set of conceptual modeling "principles" we derived in analyzing Chinese character construction [8]:

Principle I: Physical Resemblance Principle

“A hieroglyph may depict the physical shape or major visual features of the “thing” it tries to represent.”

For example, Fig. 1 has ten hieroglyphs constructed based on this principle. Fig. 1(a) shows the hieroglyph of the "lower arm,” which depicts the lower arm including the elbow and the hand." Fig. 1(b) shows the hieroglyph of a "mouth" of a person or

an animal. Fig. 1(c), 1(d), and 1(e) show hieroglyphs for "viper," "owl," and "sieve." Fig. 1(f) and 1(g) shows the hieroglyphs for “man,” and “woman.”

Fig. 1(h) shows the hieroglyph for "sun," in which the dot in the middle of the circle may represent the sunspot. It is interesting to note that this ancient Egyptian hieroglyph for sun is the same as an early form of the Chinese character for sun [8]. Fig. 1(i) shows the hieroglyph for "house." Fig. 1(j) shows the hieroglyph for "water," where the hieroglyph depicts the waves at the water surface. This hieroglyph is very similar to the early form of the Chinese character for "water," even though we did not discuss that particular Chinese character in our previous work,

All these examples show that some hieroglyphs resemble very closely the "things" they try to represent.

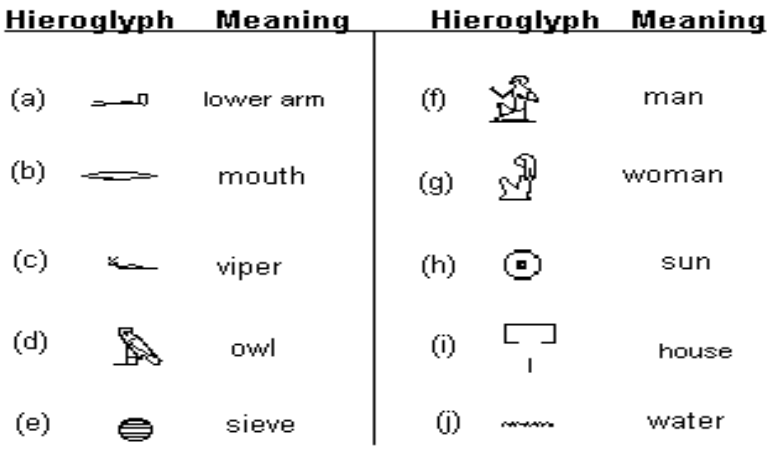

Fig. 1. Hieroglyphs based on the Physical Resemblance Principle

Principle II: The Subset Principle

“Concatenation of a hieroglyph to another hieroglyph may create a new hieroglyph which represents a subset of the things represented by the original hieroglyph.”

Fig. 2 (a) shows the hieroglyph for "house." Figure 2 (b) shows the hieroglyph for "wood." When the second hieroglyph is concatenated to the first hieroglyph, these two hieroglyphs together (as shown in Fig. 2(c)) represent a "wooden house." In this case, adding the 2nd hieroglyph may imply that a restriction (or selection criterion) is imposed to the things represented by the first hieroglyph to get a subset of them.

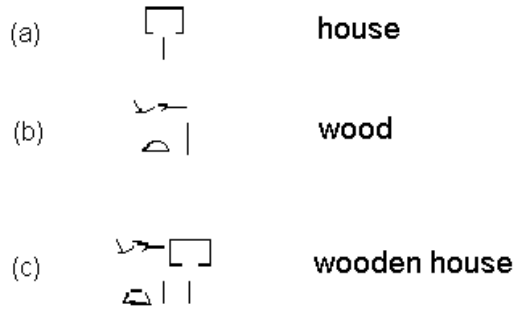

Fig. 2. Hieroglyphs based on the Subset Principle.

Principle III. Grouping Principle

"Duplication or triplication of the same hieroglyph/character has a new meaning." In ancient Egyptian hieroglyph, the new meaning usually means the "dual" or "plural" of the things of the original hieroglyph. Therefore, we can restate this principle as follows:

"A hieroglyph, which is duplex of an existing hieroglyph, may have the meaning of "dual" of the same thing represented by the original (single) hieroglyph. Along the same line, if a hieroglyph contains the triplex of another hieroglyph, it means the many instances of the things represented by the original single hieroglyph. To reduce the writing effort, the duplication (or triplication) of the whole original hieroglyph can be simplified by just duplication (or triplication) of the ending signs. In many cases, using two or three strokes at the end of the hieroglyph can also represent the duplication or triplication of the whole original hieroglyph.

Fig. 3 (a) illustrates a "god" and many "gods." The "gods" hieroglyph is a triplication of the hieroglyph of the "god" hieroglyph. Figures 3 (b) and 3(c) show two different forms of "two sisters" and "many sisters."

It is worthwhile to note that the use of this grouping principle is significantly different in ancient Egyptian hieroglyphs and in Chinese characters. In Chinese character construction, this principle is used almost solely for the generation of the new meaning of a new character rather than the pure indication of the "dual" and the "plural" concepts. Chinese language uses cardinal number "2" and the word "many" in front of the targeted word (which represents the "things") to represent the concept of "dual" and "plural," respectively. The way that the Chinese language represents the concept of "dual" and "plural" is the same as the English even though the Chinese language does not have nouns in the plural form.

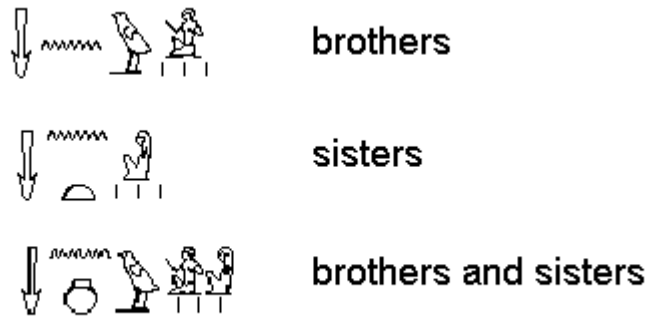

Fig. 3. Hieroglyphs based on the Grouping Principle.

Principle IV: Composition (Aggregation) Principle

"The meaning of a new hieroglyph is the combination of the meaning of its component hieroglyphs." Sometimes, it is not necessary to combine the whole hieroglyphs and only need to combine the parts of the hieroglyphs that are different.

Fig 4 depicts an example of this principle. The hieroglyph "brothers" and the hieroglyph "sisters" combine into a new hieroglyph with the meaning of "brothers and sisters."

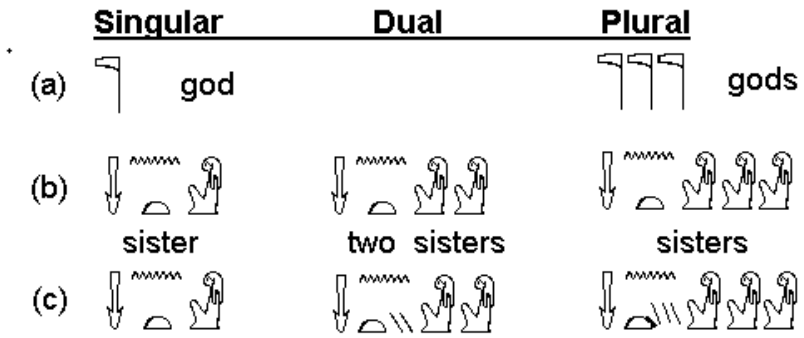

Fig. 4. Hieroglyphs based on the Composition Principle.

Principle V: Commonality Principle

"A new hieroglyph is formed by concatenation of two or more hieroglyphs. Its meaning is the common property of these component hieroglyphs." For example, in Chinese, you can construct a new character by concatenation of the character "sun" and character "moon." What does "sun" and "moon" have in common? The answer is "bright" or "brightness by light." This principle may also be applicable to ancient Egyptian hieroglyph, but we have not found a suitable example yet.

Principles VI: An-Instance-of Principle

"A component of the hieroglyph indicates which entity type it belongs to." This principle is used very often in Chinese character construction and interpretation. Usually, one part (most likely, the left part) of a Chinese character represents the (entity) type, while the other component (the right part) of the character indicates a special instance of this entity type." So far, we have no conclusive evidence that ancient Egyptian hieroglyphs also follow this principle. Further investigation is needed.

6 Conclusions and Future Research Directions

From the three languages we have studied so far, we can see the close mappings between the three languages (English, Chinese, and ancient Egyptian language) and the ER diagram technique, in particular, and the conceptual modeling techniques, in general.

Specifically, the principles used in constructing ancient Egyptian hieroglyphs (and Chinese characters and English grammar sentence structure) and the general principles used in conceptual modeling are very similar.

This is a relatively new research area. There are several ways to extend or apply this research work:

- Another natural language: One direction is to investigate other natural languages to see whether we can find something similar or different than English, Chinese, and ancient Egyptian language. A possible candidate we have in mind is the Mayan hieroglyphs.
- Better modeling and representation techniques: The conceptual modeling people can learn from natural languages (ancient Egyptian language, Chinese, English, etc.) to improve the modeling methodologies and tools.
- Improve natural language understanding and translation techniques. The people who are in the process of learning Chinese/English or building translators from one language to another can utilize some of the conceptual modeling techniques such as ERD to increase the learning speed or to design a better translator. Let us elaborate more on the point of developing new methods of learning natural language characters and statements quickly. In the past, the Chinese characters are always the roadblocks of people trying to learn more about the Chinese

written language and culture. Even though some people have advocated the use of genealogy for the learning of Chinese characters, we think there is a very promising direction and much more work is needed.

- Better understanding on how a human thinks: It should be interesting to investigate how human mind works in recognition of the hieroglyph set either with or without the help of the construction principles discussed in this paper.
- Better interfaces between natural language interface and DBMS/repositories: Since ERD is used heavily in DBMS and repositories, the user interface design for these system could take advantage of the correspondence between ERD and natural languages.
- Better Design of GUI Icons: The icons in the graphical user interfaces of many systems or application software packages not correctly designed. As a result, the meaning of the icons confuses a lot of users. It will be nice to use the result of this work as a foundation to build a methodology for icon design.
- Using cognitive type & cultural background to select the "better" conceptual modeler: In a team environment, it is important to put the right person to do the right job. Not everyone is a master of natural languages. Perhaps those who can master natural language usage would be good candidates for the job of conceptual modeling.
- Using "gender" as a one of the factors for selection of "better" data/knowledge modeler: Are women more qualified than men to do conceptual modeling? This is a politically sensitive issue, but it may merit a scientific investigation.
- Adopt modeling methodologies and techniques to local environments: Use cognitive type and cultural background to select the most natural modeling techniques and methodologies for a particular environment.

References

1. Embley, D. and Goldstein (ed.), *ER'97*, LNCS, Springer, Berlin, 1997.
2. P. P. Chen, The entity-relationship model: toward a unified view of data, *ACM TODS* 1 (1976) 1-36.
3. P. P. Chen, English sentence structures and entity-relationship diagrams, *Information Sciences*, (1983) 127-149.
4. P. P. Chen, The time-dimension in the entity-relationship model, in: H..-J. Kugler ed., *Information Processing* (Amsterdam, 1986) 387-390.
5. P. P. Chen and A. Zvieli, Entity-relationship modeling of fuzzy data, *Proc. of* 2nd *International Conf. on Data Eng.* (Los Angeles, 1987) 320-327.
6. P. P. Chen, The denotational semantics of the entity-relationship model, (with N. Chandrasekaran and S.S. Iyengar), *International Journal of Computer Mathematics* (1988) 1-15.
7. P. P. Chen, ER modeling for multimedia applications on the internet, *Proceedings of 1995 Conference on Applications of Databases, ADB '95*, San Jose, CA, 1995 (keynote speech, abstract only).
8. Chen, P. P., English, Chinese and ER Diagram, *Data & Knowledge Engineering, Vol. 23*, No. 1 (June 1997), 5-16.

9. Egyptologica Vlaanderen VZW, *Reading Hieroglyphs: The First Steps, Lessons 1-8, http://hosting.netvision.be/egyptologica /e_home.html, 1998.*
10. Gardiner, Sir Alan H., *Egyptian Grammar*, Griffith Institute, Oxford, 1979.
11. Harbaugh, R.(ed.), *Chinese Characters: A Genealogy and Dictionary*, Zhongwen.com., 1998.
12. Loucopoulos, P. (ed.), *ER '94*, LNCS 881, Springer, Berlin, 1994.
13. Papazoglou, M. P. (ed.), *OOER '95*, LNCS 1021, Springer, Berlin, 1995.
14. Ling, T. W., Ram, S. and Lee, M. L.(eds.), *Conceptual Modeling -- ER '98*, LNCS 1507, Springer, Berlin, 1998.
15. Laplante, P.(ed.), *Great Papers in Computer Science*, West Publishing Co., St. Paul, Minnesota, 1996.
16. Pestman, P.W. (ed.), *Aspects of Demotic Lexicography*, Leiden, 1987.
17. *The Software Challenges*, in the series: Understanding Computers, Time-Life Books, 1993, Third printing.
18. Thalheim, B. (ed.), *ER '96*, LNCS 1157, Springer, Berlin, 1996.
19. Thalheim, B., *Fundamentals of Entity-Relationship Models*, Springer, Berlin, 1997.
20. Whitten, J. L. and Bentley, L. D., *Systems Analysis and Design Methods*, Irwin/McGraw-Hill, 1998, 4th Edition.
21. Yang, A. and P. P. Chen, Efficient data retrieval and manipulation using Boolean entity lattice, *Data & Knowledge Eng.*, *Vol. 20*, No. 2 (October 1996), 211-226.