Archeologia e Calcolatori
7, 1996, 703-711

A "TREETEXT" AND PICTURES MANAGER AS A PERSONAL TOOL FOR OBJECT-KNOWLEDGE REPRESENTATION IN CLASSICAL ARCHAEOLOGY AND HISTORY OF ART

1. Description

In archaeology and history of art the distinction between object-based and method-based knowledge is useful. The first type refers to knowledge about the constitution and the qualities of the single real objects of these sciences; the second type is the knowledge won by the comparison of such objects [1]. At first glance the latter one may seem to be of more interest with regard to the historical evaluation of archaeological artefacts, with object-knowledge remaining only a trivial question.

But a closer examination of the usual objects to be dealt with in archaeology and history of art, as for instance Greek or Roman temples, medieval churches, or works of figurative arts (such as Greek vases or Greek and Roman statues and reliefs or Roman wall paintings), reveals their characteristically complex structure. The comparableness of such objects very often is limited to the detail level, such as capitals in architecture or iconographic details in figurative arts. Repetitions of complex objects which would allow one to establish typologies at a high level are relatively uncommon.

It is apparent that the basic requirement which a useful tool for working with object-knowledge must fulfill is the capability to handle the textual representation of objects which have a high degree of complexity. For the tool which I will subsequently be describing it was first necessary to develop the description language "TreeTexT" and afterwards a parser and a retrieval module both able to manage also pictures of such objects.

Object descriptions normally proceed from the whole to the parts, subparts and so on. The relation between the components and their respective subcomponents is called the "partitive relation" (Neveling *et al.* 1975, 100 see nr. 41-42-4). A common method of representing this relationship is the tree structure, where the root marks the whole and the different branches down to the leaves reflect the organization of the respective part and subpart

[1] Djindjian 1993, Moscati 1990b, Moscati 1994, Orlandi 1993 treat and explain both types of knowledge; Guimier-Sorbets 1990 is more oriented to the object-based or "documentary" approach and Guermandi *et al.* 1994 and Modrzewska *et al.* 1993 and Moscati 1990a to the comparative approach using quantitative methods. A third approach proposed by Gardin *et al.* 1987 and Lagrange *et al.* 1994, the "inferencial" one, works with methods used in artificial intelligence known as "rules" and should allow object-based as well as method-based applications. For all approaches refer to the basic work of Moscati 1987 and for an integrated model using the relatively new "object-oriented" approach refer to Missikoff 1995.

dependencies (e.g. GUIMIER-SORBETS 1978, 130 and more detailed GUIMIER-SORBETS 1990, 143-182) (Fig. 1).

A formulated "TreeTexT" like in one of my databases about Greek architecture (Fig. 2) normally consists of different descriptions or "documents" about complex objects. The tree-structured organization of a single document is represented by "contextors" of one or more lines, which are the main division marks inside a document, separating the text into sections belonging to the single parts of an object. Contextors may be understood as a simple punctuation. The sections of text between contextors contain assertions about the corresponding parts (or contexts) of the described object. Normally such an assertion consists of one or more keywords marked by an asterisk and called "descriptors". Words without an asterisk are limited to the role of commentaries. Descriptors may belong to the textual type used for names and denominations or to a numeric type allowing the representation of named integer or real values or ranges.

Textual descriptions of objects can be more or less explicit. The description of an ancient temple may be superficial and generic or deep and detailed. For scientific purposes the latter case is undoubtedly preferable. However many of the real objects of classical archaeology are more or less fragmentary which means that often only a very limited description is possible. The fundamental idea of "peripteros" and its different appearances as pentastyloi to enneastyloi having 10 to 17 columns in depth (Fig. 3) are connected by the "generic relation" (NEVELING *et al.* 1975, 99 see nr. 41-41-4); this relationship [2]can be represented with a tree structure going from abstract to concrete (for instance GUIMIER-SORBETS 1978, 68). The same relation exists between "octastylos" (and "octastylos: 17", a mere thesaurus-pointer which is allowed to be used only in retrieval mode but not in a description) and "octastylos: 17: peript" for the peripteroi and "octastylos: 17: psdipt" for the pseudo-dipteroi. Another similar case is given by the goddess "Athena" and her different regional personifications like "Athena: Alea", "Athena: Aphaia", "Athena: Parthenos" or "Athena: Polias" (the last one with her subpersonification "Athena: Polias: Nikephoros").

Establishing such generic links is the main function of a "thesaurus" and the first document of a "TreeTexT" should be used for this purpose; a further usage of the thesaurus is to provide terminological control of the

[2] The recent approach to documentary data by MISSIKOFF 1995 relies mostly on the use of the generic relation. According to a paradigm introduced in the eighties by software developers, this approach is called "object-oriented" and is based upon hierarchically ordered "classes" ranging from an abstract (or generic) root to increasingly concrete subclasses (MISSIKOFF 1995, 236-239 demonstrates this stepwise procedure applied to different fibula types). A class contains not only a data structure describing the single objects or "instances" but also specific "methods" or programs for handling them. (In the field of archaeology that could be for example methods for specific retrieval or for statistical evaluation). Methods developed for a given abstract object level are "inheritable" or to be applied also to a more concrete level, but not vice versa.

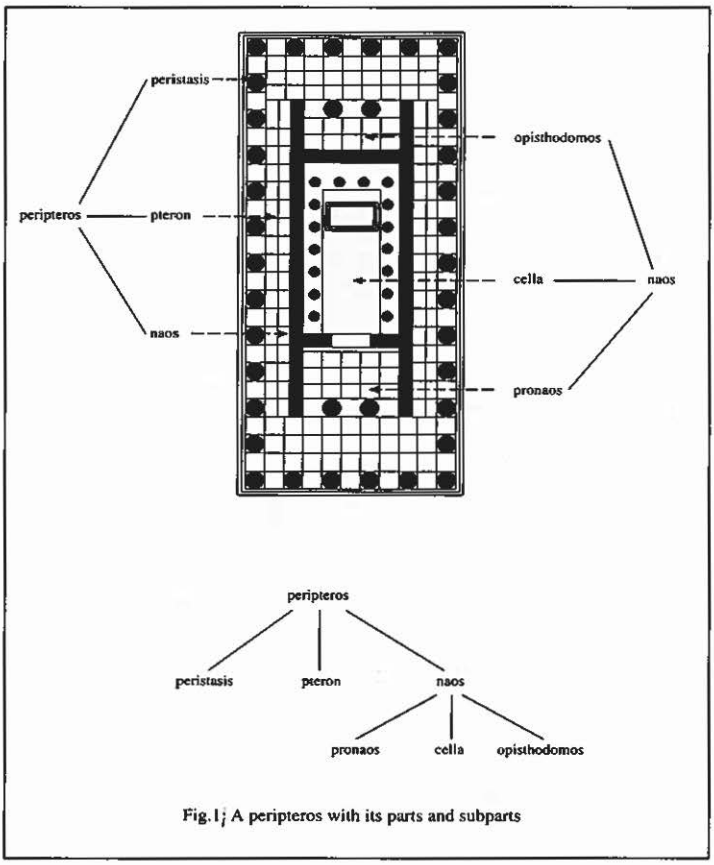

Fig. 1 – A peripteros with its parts and subparts.

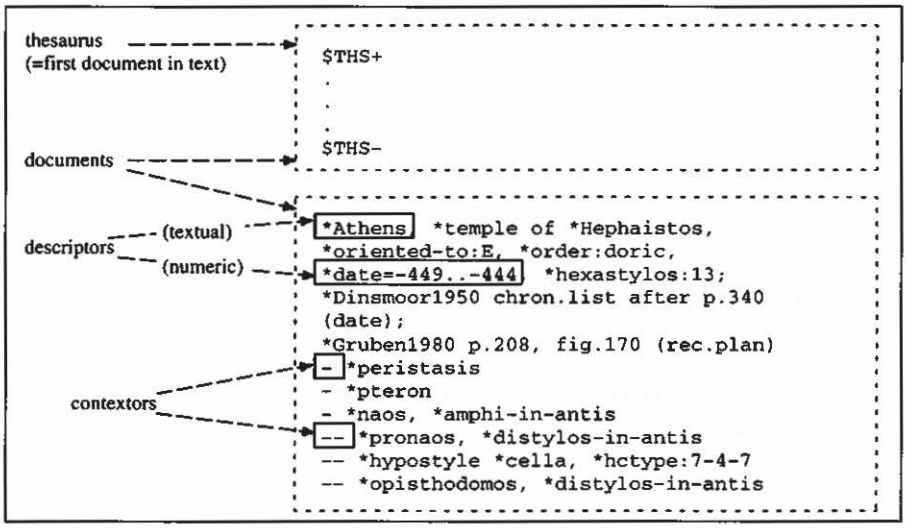

Fig. 2 – "TreeText" with a peripteros description.

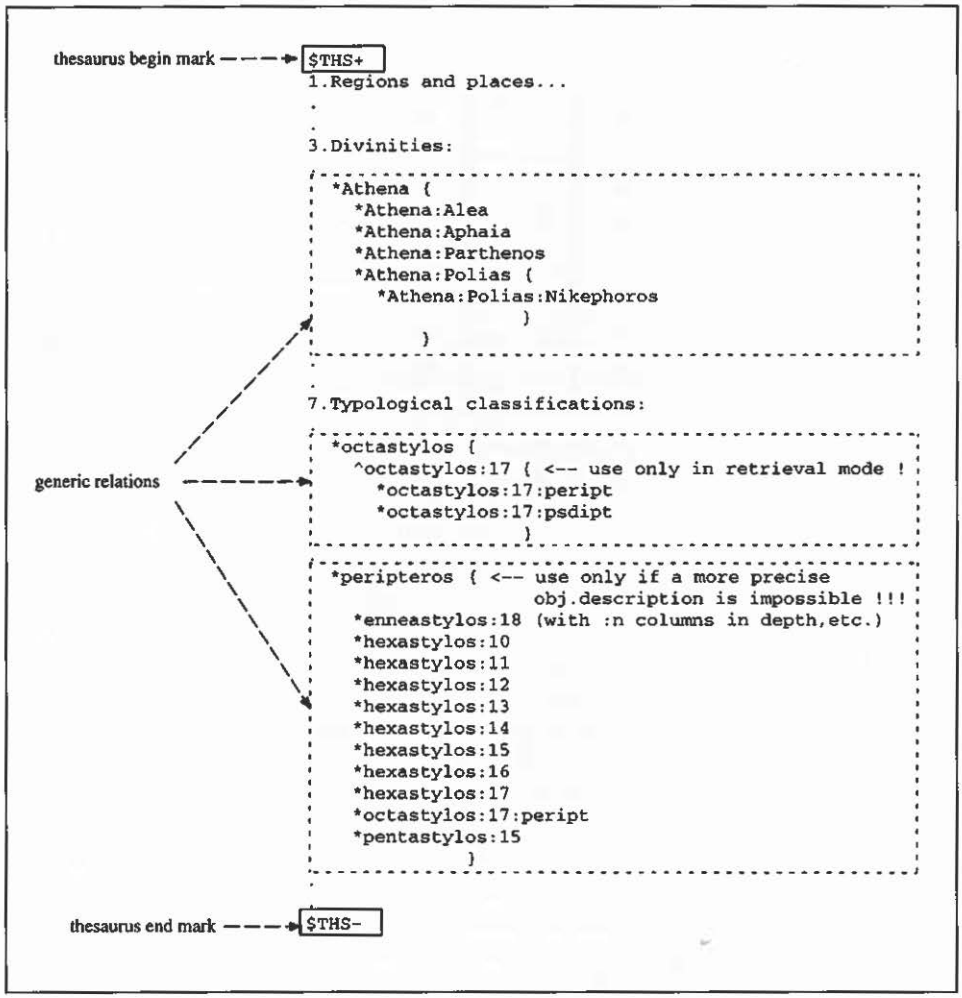

Fig. 3 – Description of generic relations in a "TreeText" thesaurus.

descriptors in the following documents. The first document must be delimited by special thesaurus marks and cannot contain contextors. Generic relations are expressed by enclosing the logical subordinate descriptor set in braces. Nested relations are possible.

Experiments with students (at the Universities of Frankfurt/M.,Graz (Austria) and Hamburg in the years between 1986 and 1990) as users of an earlier version of the "TreeTexT" data managers (the "ARBOR" managers, see EISNER 1989, EISNER 1990) showed that a mere textual description and documentation of archaeological objects is not satisfactory; there always re-

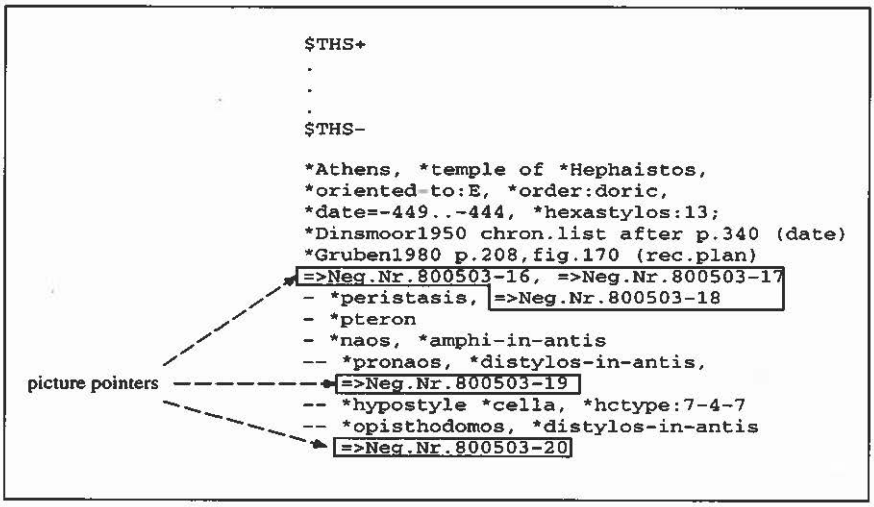

Fig. 4 – Picture pointers in a "TreeText" description.

mains a need for references to illustrations or pictures like usual in scientific literature. So a new type of flag has been introduced in the description language, the "picture pointer" which appears as an arrow (Fig. 4). Every context can contain as many of these pointers as necessary and wanted for illustration of the respective document part. But the relation established between the text and the picture set is only a logical one and not of layout type[3].

A syntactically correct "TreeTexT" description file passes through a parser which generates a representation of the data in a form which allows a human user to search for textual and pictorial information about real objects. Retrieval is clearly text-based, i.e. upon descriptors which can be combined in very different ways reflecting the hierarchical organization of the searched objects and of the thesaurus entries. The retrieval results (including pictures if present) may be viewed using the included browsing facilities.

2. Aims

Due to the capability of a "TreeTexT" retrieval manager to establish links via descriptors between differently structured complex objects, this tool may be used for a variety of purposes such as the production of personal scientific notebooks containing data about scientific literature and discussion themes, or the production of illustrated terminology databases for students;

[3] Contrary to most mixed applications as for instance Benelli 1992, Gudivada *et al.* 1994, the TreeTexT software package does not support multimedial documents composed of a mixture of textual and pictorial information but only textual documents and subdocuments with a variable number of pointers to the appropiate pictorial documents.

```
---------------------------------------------------------------------------
TreeTexT by Eisner: Rtrv.Vrsn.X.1.6 [95-09-20]; 95-09-26 11:02:26
---------------------------------------------------------------------------
The Actual Query Status Is:
  (0) {THS-}
? naos
  Hits In 29 Documents
  (1) {THS-} [Searching In Sub-Contexts]
? staircase=2
  Hits In 6 Documents
  (2) {THS-} [Searching In Primary Context]
? prostylos
  Hit(s) In 1 Document
---------------------------------------------------------------------------
*Paestum, *temple of *Athena ("Demeter"), *oriented-to:E, *order:doric,
*date=-515..-505, *hexastylos:13
*Dinsmoor1950 chron.list after p.340 (date);
*Gruben1980 p.251,fig.200 (rec.plan)
- *crepidoma:2
- *peristasis
- *pteron
- *naos, *prostylos
-- *pronaos, "*prostylos-cum-antis", *tetrastylos:2
--- *anta, *pseudo-column
-- *cella
-- *staircase=2
```

Fig. 5 – Result obtained after a primary query for the term "naos", for "staircase = 2" in sub-contexts and for "prostylos" in the primary context.

```
---------------------------------------------------------------------------
TreeTexT by Eisner: Rtrv.Vrsn.X.1.6 [95-09-20]; 95-09-26 12:58:06
---------------------------------------------------------------------------
The Actual Query Status Is:
  (0) {THS-}
? cella
  Hits In 26 Documents
  (1) {THS-} [Searching In Adjacent Contexts]
? cella
  Hit(s) In 1 Document
---------------------------------------------------------------------------
*Corinth, *temple of *Apollo, *oriented-to:E, *order:doric, *date=-545..-535,
*hexastylos:15
*Dinsmoor1950 chron.list after p.340 (date);
*Gruben1980 p.100,fig.95 (rec.plan)
- *crepidoma:3
- *peristasis
- *pteron
- *naos, *amphi-in-antis
-- *pronaos, *distylos-in-antis
-- *hypostyle *cella 1, *hctype:4-0-4
-- *hypostyle *cella 2 (accessible from opisthodomos !), *hctype:2-0-2
-- *opisthodomos, *distylos-in-antis
---------------------------------------------------------------------------
```

Fig. 6 – Result obtained after a primary query for the term "cella" and for "cella" in adjacent contexts (i.e. for objects with more than one cella).

alternatively it could be used as a management tool for maintainig photographical archives.

The production of a CD-ROM base system with an initial database which may subsequently modified and enlarged according to personal requirements is a conceivable practical application.

3. Results

The actual implementation of the "TreeTexT" parsing and retrieval modules is based upon earlier versions of ARBOR which was originally developed under DOS with Modula-2. For easier portability and interfacing with existing software, for example image-processing or database modules, in this implementation C and C++ are used as programming languages and Solaris and Linux as operating systems on Sun Workstations and IBM-compatible PCs.

```
-------------------------------------------------------------------------------
TreeTexT by Eisner: Rtrv.Vrsn.X.1.6 [95-09-20]; 95-09-26 12:45:37
-------------------------------------------------------------------------------
The Actual Query Status Is:
  (0) {THS+}
? naos
  Hits In 30 Documents (Including Thesaurus)
  (1) {THS+} [Searching In Primary Context]
? in-antis
  Hits In 25 Documents (Including Thesaurus)
  (2) {THS+} [Searching In Root Context]
? date=-550..-450
  Hits In 11 Documents
  (3) {THS+} [Searching In Next Sub-Contexts]
? tetrastylos
  Hit(s) In 1 Document
-------------------------------------------------------------------------------
*Selinus, great sanctuary, *temple 'G' of *Apollo, *oriented-to:E, *order:doric,
*date=-520..-450, *pseudo-dipteros, *octastylos:17
*Dinsmoor1950 p.99 (Apollo), chron.list after p.340 (date);
*Mertens1984 P.164,fig.78 (rec.plan)
- *crepidoma:2
- *peristasis
- *pteron
- *naos, *prostylos-et-in-antis
-- *pronaos, **prostylos-cum-antis", *tetrastylos:2
--- *anta, *pilaster
-- **hypostyle" *pseudo-cella, *sekos, *hctype:10-0-10
--- *adyton
-- *opisthodomos, *distylos-in-antis
-------------------------------------------------------------------------------
```

Fig. 7 – Result obtained after a primary query for the term "naos", for "in-antis" in the primary context, for "date=550..-450" in the root context and for "tetrasytlos" in the next sub-contexts relative to the primary one. Due to the use of the flag "{THS+}" the query term "in-antis" led to the additional descriptor "prostylos-et-in-antis" and "tetrastylos" to "tetrastylos:2".

At the moment of writing, the GUI of the retrieval module, which is based on OSF/Motif, is provisionally working but not completely finished. A simple picture server still needs to be found or developed and different additional possibilities of browsing through pictures remain to be evaluated and tested. A database with some thousands of documents referring to my personal photographic archive (Greek and Roman objects in the Classical World) and with bibliographical data serve as experimental material.

In order to get a first impression of how the "TreeTexT" retrieval manager works please refer to the attached printed output which is produced as a result of some queries to the mentioned database about Greek architecture (Fig. 5 to 7). These illustrations provide examples of the use of various operators (e.g. the operator for searching in adjacent contexts, subcontexts or in the root context) which allow the stepwise refinement of the results when searching in hierarchically organized "TreeTexT" documents. The flags "{THS+}" and "{THS-}" indicate whether or not the special thesaurus functions were working during the single retrieval step. Output in picture form has not been included due to layout considerations.

MICHAEL EISNER
Darmstadt

Acknowledgements

I am grateful to Natalie Böhm and to Hannelore Eisner for their support of my work.

BIBLIOGRAPHY

BENELLI M.V. 1992, *Le collezioni archeologiche dell'Opera della Primaziale Pisana. Un sistema ipertestuale per la catalogazione automatica di materiali antichi reimpiegati*, «Archeologia e Calcolatori», 3, 139-155.

DJINDJIAN F. 1993, *Les systèmes d'informations en archéologie*, «Archeologia e Calcolatori», 4, 9-25.

EISNER M. 1988, *ARBOR. Un approccio sistematico ai reperti di struttura complessa*, in A.M. LIBERATI SILVERIO (ed.), *Archeologia e informatica. Analisi delle metodologie e delle tecnologie dell'informazione in archeologia con particolare riferimento alla cartografia, allo scavo, alla catalogazione ed agli aspetti museali e didattici*, Roma, Quasar, 21-26.

EISNER M. 1989, *ARBOR. Eine Sprache zur Beschreibung und ein Programmpaket zur Verarbeitung hierarchischer Datenobjekte der klassischen Archäologie und Kunstgeschichte*, Birlinghoven, Gesellschaft für Mathematik und Datenverarbeitung mbH, GMD-Studien 159.

EISNER M. 1990, *The ARBOR information system for classical archaeology and history of art*, «Archeologia e Calcolatori», 1, 253-262.

GARDIN J.-C., GUILLAUME O., HERMAN P.-Q., HESNARD A., LAGRANGE M.-S., RENAUD M., ZADORA-RIO E. 1987, *Systèmes experts et sciences humaines. Le cas de l'archéologie*, Paris, Eyrolles.

GUDIVADA V.N., RAGHAVAN V.V. 1994, *A System for Retrieving Images by Content*, «RIAO 1994 Papers», 418-436.

GUERMANDI M.P., MIGNANI S., MONTANARI A. 1994, *Il trattamento dei dati iconografici nell'ambito della ceramica attica e l'esempio del progetto Spina,* «Archeologia e Calcolatori», 5, 111-140.

GUIMIER-SORBETS A.-M. 1978, *La constitution des données en archéologie classique*, Paris, CNRS.

GUIMIER-SORBETS A.-M. 1990, *Les bases de données en archéologie. Conception et mise en oeuvre*, Paris, CNRS.

LAGRANGE M.-S., CONCEIÇAO RODRIGUES M. da 1994, *La représentation d'un raisonnement en archéologie. Un exemple: analyse logiciste et système expert,* «Archeologia e Calcolatori», 5, 333-353.

MISSIKOFF O. 1995, *Proposta di applicazione di un approccio object-oriented alla formalizzazione di dati qualitativi,* «Archeologia e Calcolatori», 6, 1995, 225-242.

MODRZEWSKA I., TARONI G. 1992, *Analisi quantitativa di forme di alcune lucerne tardoromane,* «Archeologia e Calcolatori», 3, 19-36.

MOSCATI P. 1987, *Archeologia e calcolatori*, Firenze, Giunti.

MOSCATI P. 1990a, *Indirizzi e sviluppi dell'archeologia quantitativa*, in P. MOSCATI (ed.), *Trattamento di dati negli studi archeologici e storici*, Roma, Bulzoni, 1-54.

MOSCATI P. 1990b, *L'analisi quantitativa nell'archeologia di epoca storica,* «Archeologia e Calcolatori», 1, 39-80.

MOSCATI P. 1994, *Un gruppo di urne etrusche di produzione volterrana: prospettive di analisi quantitativa,* «Archeologia e Calcolatori», 5, 87-110.

NEVELING U., WERSIG G. (edd.) 1975, *Terminologie der Information und Dokumentation*, München, Dokumentation, DGD-Schriftenreihe 4.

ORLANDI T. 1993, *Sulla codifica delle fonti archeologiche,* «Archeologia e Calcolatori», 4, 27-38.

ABSTRACT

"TreeTexT" is a formal language developed for describing the real objects (or monuments) of classical archaeology and history of art. It works with the "partitive" relation (i.e. between the whole, its parts and subparts) in order to permit the handling of complex objects such as Greek or Roman temples. A "thesaurus" allows to establish terminologies based upon the "generic" relation (i.e. between more or less explicit descriptors). A software package manages not only a specific textual retrieval at every described detail level of complex objects but also possibly related pictorial information.