Sites & Monuments
National Archaeological Records

Printed in Denmark by Sahva Håndværkskolen Bogtryk/Offset 1992
Composition: The National Museum of Denmark, Documentationdept.
DTP: Aldus Page Maker
Fount: Palatino 10/12
Paper: 115 gram Uno-matt

Edited by Carsten U. Larsen
Designed by Elisabeth Halgreen og Hanne Carden

ISBN 87-89364-02-3

Cover: Dolmen from NW-Zealand, Denmark, 1880's A.P. Madsen del.

Contents

Data Structures for Excavation Recording A Case of complex Information Management

Jens Andresen and Torsten Madsen

Many archaeologists, especially in England and the United States, to a lesser degree in continental Europe, and hardly at all in Scandinavia, use computers to handle excavation data. This is done either directly in the field, and all the way through to publication, or exclusively in the post excavation process.

Archaeological excavation, is a well defined methodological and technical area within archaeology. As a consequence, data management systems designed to handle excavation data should be expected to be quite similar in structure and function, but they are not. There seems to be a marked diversity in the systems present. One reason for this is that even if archaeological excavations are methodologically and technically well defined, there are nevertheless major differences in the way that excavations are carried out, not least with respect to the types of excavated sites (e.g. deeply stratified town sites *versus* unstratified rural sites). Regardless of the type of site, however, archaeologists also have widely varying ideas and preferences where documentation of excavations is concerned. These differences are noticeable at all levels - between countries, between institutions within countries and even between persons within institutions. The magnitude of these differences is difficult to assess, however, because amazingly few contributions concerning the structure of excavation data have been published.

First of all one may point to Harris' classic work (1979), which more than anything else has served as the foundation for attempts to improve recording systems for archaeological excavations. These attempts seems to have developed in two directions. One is the "single context planning" scheme advocated by the Department of Urban Archaeology at the Museum of London (Spence 1990). "Contexts" here comprise all types of observational units viewed as being equally important. The other is a strictly hierarchical system advocated by Carver (1985) and used by the Central Excavation Unit (Hinchliffe & Jefferies 1985) and the Archéo-Data system (Arroyo-Bishop 1989, 1991) to mention a few published systems. The latter approach is perhaps the most influential, with its appealing hierarchical structure, which can be delineated with a very brief citation from Carver: "The assumption made here is that archaeological data are susceptible to unequivocal definition within a conceptual hierarchy comprising *component*, *context*, *feature* and *structure*, each of which consists of a set of the others" (1985 p. 50).

In Denmark it is also by now customary to use a hierarchical system not very different from that of Carver. Its lowest level is *fund* (= component), the next level is *lag* (= context), and the upper level is *anlæg* (= feature and structure). Many archaeologists consider *anlæg* to be the primary concept, and as a consequence an excavation is often recorded in terms of *anlæg* described by other *anlæg* and/or by *lag*. In some cases, mostly in connection with specific types of sites, where the separation of *anlæg* is far from clear, you often find the excavation recorded entirely by means of *lag*.

The diversity in recording practices leads to another reason why existing computer systems appear so different. Whenever an information management system is to be established, it is customary to make an investigation of what actions the users of the system *in spe* perform in their daily work, what types of procedures the system should comprise, what output the system is going to produce, and what kind of user-interface the system should present. Based on this investigation the designer of the system will decide the basic structure, and will most often do this in the simplest possible way, strongly taking into consideration what the chosen database management system will allow.

One problem, or perhaps rather nuisance, with a system created this way, is that it often ends up being a one man or at least one institution system. As the system reflects the preferences and habits of the creators, there is a tendency for other archaeologists to find it unsuitable for their needs, as it does not reflect their own ways of recording.

Another problem with this procedure, leading in our opinion to a regular flaw, is that by direct transformation of a traditional registration system into a computer information system, you not only transfer the layout of the hand written system to the computer, you also transfer the logic of written records to the computer media. This of course is not a failure by itself, as you have exactly the same presentation on computer that you had in hand writing, but you do not take full advantage of the change in media. The questions not being asked in these direct transformations are, whether the traditional hand written recording systems do full justice to the informations recorded, and whether the computer media is able to handle the structure of excavation data in a better way than is possible with the flat paper file systems.

In order to create an information management system for archaeological excavations it is imperative, we would argue, to start out with a discussion of the logic of the information to be handled. Before we can discuss what information we would like to maintain, it is necessary to discuss the structure of information pertaining to archaeological excavations in general, and to do so from the point of view that the medium on which to structure our data is a computer and not paper. What are the basic types of information that we handle? What is the internal logical structure of these types? and what kind of relations exist between the various types of information?

The procedure, where you start by analysing the user activity pattern may in many cases be an excellent way into database design, but it seems inappropriate in cases where - as in archaeology - the structuring of the data is open to theoretical discussions, and where indeed the structuring affects the results obtained (Ryan & Smith 1986: 56). In other words we wish to start out by questioning the concepts that archaeologists hold on the nature of the archaeological data they document in the field.

As mentioned above the indisputably most dominating organizing concept is hierarchy. One could, as Carver does, claim that this is due to the nature of archaeological data, but in our opinion it is more due to the appealing simplicity of representation that this offers. Hierarchical organization of information is very easy to comprehend and it appears very clear when presented on paper. In fact it is *the* accepted way of presenting complex information on paper as any scientific publication with its chapters, sub-chapters, sub-sub-chapters etc. can show. It may, however, not be the best way to represent the totality of complex information, as the heavy need for indexes in scientific publications tend to demonstrate.

In Carver's hierarchical universe a potsherd (= component) could stem from a layer (= context), constituting the fill of a posthole (= feature) that forms part of a house (= structure). This would be a very straight forward way to describe an actual set of observations in an excavation. But what if the posthole was part of one end wall of the house, and this wall was integrated into a fence surrounding the farmstead? Or what if the potsherd fitted together with a potsherd from a layer in another posthole? This information would not fit the hierarchy, and it would be necessary to add additional cross references to represent them. The problem with hierarchies is that they represent a particular, univariate instantation of our view of archaeological data, and not the multivariate reality of how we view the data.

Further it is a problem with a hierarchy like Carvers that it has a fixed number of levels, where we as archaeologists would like to make abstractions in the indefinite (for instance, house - farm - village - parish for the example mentioned here) reflecting the way we perceive the actual archaeological situation we are dealing with.

In the following we shall try to give a rough outline of a logical design of an information system for archaeological excavations. The design is based on relational data management theory and is intended to overcome the shortcomings of hierarchical organisation. In addition to the outline of the design we shall present a minor exemplification to make the concepts of the design more understandable.

The properties of excavation data entities

Stratigraphical context informations - the layer concept

It is a stated fact that the concept of stratigraphy and the formulation of its basic principles comes from geology. Whether the rules of stratigraphy, however, are the same for geology and archaeology is more debatable. Harris holds the opinion that the stratigraphic rules in archaeology are more complex and comprehensive than in geology, since human beings are capable of creating much more complex stratigraphical phenomena than is nature.

Harris divides stratification into strata and interfaces. Strata (also called layers or deposits) are further split into natural strata and man-made layers, the latter with the subtype upstanding strata. The difference between natural strata and man-made layers is only that the latter is planned and laid down by human beings exactly as we find it, whereas the former has an element of natural deposition associated (erosion of some sort). Interfaces are divided into layer interfaces, feature interfaces and period interfaces. Two forms of layer interfaces are defined: surfaces of strata which have been deposited in a more or less horizontal manner, and the vertical surfaces of upstanding strata. The feature interfaces are formed by the destruction of pre-existing stratification, and are split into two types, the horizontal (demolition interfaces on walls for instance) and the vertical (pits, postholes etc.). The period interface represents a major break in a stratigraphical sequence, where the formation of layers between interfaces can be separated by cultural, chronological or functional criteria.

It is obvious that Harris classified archaeological deposits based on interpretations of their origin, mechanisms of deposition, and post deposition processes. As his classification of strata is based not on described, physical qualities, but on assumed, historical properties, he ends up with a system where it is difficult to separate between description and interpretation. Thus the distinction between natural strata and man-made layers is imbedded in the description of these, even if it must be considered an interpretation. And further, as an interface is given a number in the layer sequence, it obtains a *de facto* position as a layer along with natural strata and man made layers, even if the interfaces in our opinion must be considered to be interpretations of the observed layering, and not a part of the layering itself.

Although Harris focussed on the division line between layers, when he speaks of interfaces, he nevertheless introduces the idea of composite structures in practice. A feature interface may in Harris' term be the surface of a cutting for a pit, but to be of any practical use it has to be considered as the delimitation of the layers found within the pit. The layers "belong" to the pit. By the same token a period interface is of little use as such. What is of interest is the assemblages of layers between the interfaces.

Harris does not operate with structures, but others do as mentioned in the introduction, and this complicates the issue further. Structures are normally viewed as a combination of layers and interfaces that forms a conceptual totality. A house for instance may be defined as a combination of feature interfaces (postholes and wall trenches), and man-made layers (a clay-laid fireplace for instance). The latter however could itself be a structure, if it consist of a combination of two man-made layers (clay layer and stone frame). It is easy to see how the ambiguity grows, when suddenly the same thing in one instance can be viewed as a layer, and in the next as a structure.

In Danish archaeology, it is customary as mentioned to operate with *lag* and *anlæg* only. The *anlæg* are considered to be the primary concept, as an excavation is divided into a number of *anlæg*, which then are described by way of physical *lag*. The *lag* are furthermore separated and numbered within *anlæg* only. This procedure may seem logical and straight forward - you observe some features/structures, and then describe them with reference to their constituent parts. The logic, however, breaks down on two points. Firstly, it is often not possible to separate *anlæg* on an excavation, either because there are no obvious *anlæg*, or because you cannot identify them within the opened area. Descriptions of the "constituent parts" then, are either done "in a vacuum" with no reference to an *anlæg*, or it is organised through the use of artificial *anlæg* like "trench", "level" or "area". Secondly, even if the idea is, that the whole of the excavation should be divided up into neat separate *anlæg*, most excavations ends up with a hierarchy of *anlæg*, where higher level *anlæg* are described as consisting of lower level *anlæg*.

The flaw in these systems, not least in an elaborate system like Harris', is that there is no clear separation between observation and interpretation. This is noted by Julie Stein (1987), who points to geology, where it is now customary to distinguish between observable units and inferential units. The former are defined by their physical properties alone, which means that they are separated through their described structure, and their delimitation to other units with a different structure, exclusively. The latter category is based on an interpretation of the observed units and their content.

We agree with Julie Stein: it is very important that there is a strict distinction between observed "natural" layers on the one hand, and inferential units on the other, in archaeology as well. The main reason is that although we may feel absolutely certain that a layer we observe in the field represent a posthole, and hence would tend to register it as such, there is a major difference between the layer as such, and the layer categorized as representing a posthole. Layers and all information relating to layers are entirely a matter of on site registration. All decisions of how a specific layer should be delimited, and what qualities should be assigned to it, must be taken during the course of excavation. The records that hold the information on observations of layers attain a status as historical documents, whenever the possibility of affirmation is obliterated, and must be treated as such; that is, they are never to be changed. It should however be noted that it is the observation "a layer was seen" that represents a historical fact. We may on second thought, even after the excavations is closed, decide that observations, which claim two different layers, were actually referring to the same layer. Indeed the current practice is to document the layers in a number of plans and sections crosscutting the deposits. The same physical layer may thus be seen in a number of plans and sections, and here be identified as different (observations of) layers.

The assertion that an observed layer represents a posthole is, on the other hand, of an entirely different nature. In the field we may be convinced that we are dealing with a posthole, and then later on change our minds. This is fully legitimate as an act of interpretation, and the records that keeps information on how observations of layers should be interpreted should be editable at any time.

The observable units and the inferential units are thus two absolutely distinct entities, and should be treated as such. In this paper we term the observable units layers, while the inferential units are termed constructs (a better word should probably be coined, if possible). These will be treated in a separate section.

The entity layer, is thus **the** physical unit of context information in archaeological excavations. It is simultaneously the largest and the smallest unit observed. Exactly how it is delimited is entirely a question of definition in the hands of the excavator. He may choose to define very coarse layers or very fine layers. As long as the layers are observable in the field, he may alter his registration by redefining a number of layers to be just one layer, and delete the extras, or he may split a layer into two or more, provided that he can handle the registrations of moveable objects already made with respect to the new layering. But he may not define a new layer as a grouping of layers, and at the same time maintain these as sub-layers to the layer. This would transgress the borderline to constructs.

Each layer defined possesses a set of physical properties that can be described, as well as stratigraphical relations to one or more other layers. It is customary to think of stratigraphical relations as something complex and manysided. However, in connection with layers defined as direct observational units it is quite simple. Basically, it is sufficient to register the relation "above" (or "below"), but as it should be possible to set two layers to be the same, an option of "same as" is also a necessity. For example layers in two neighbouring trenches may initially be given different numbers, and be referenced on several drawings. When a correspondence between the layers in the two trenches subsequently is established, it becomes unpractical to renumber them, not for the computer system, but for the drawings, and the photographs taken with the original numbers on. The key problem is really that it is not the layers we number, even if we call them layers, but our observations of the layers. A "same as" option is thus necessary in order to link different observations of the same layer. Because any layer may be found "above" N other layers, and any layer may be the "same as" N other layers, these may relate to each other in an N:M fashion, and may thus in terms of relational theory be represented by two tables, one representing the observations of layers as such, and one representing relations between layers.

If we reserve the table "relations between layers" for information concerning stratigraphical relationships, we may further note that a binary relation between two layers can only occur once, because the relation between two layers cannot be both "above" and "same as". Further we should note that every member in the relations table must have one and only one member in the layers table. In other words no relationships of the types 0:1, 0:M, 1:0 and N:0 can occur at any time. Also, this indicates that the lower layer and the upper layer in a relation always will end up in the same ID-fields of the relationship table, and further that the physical structure of a site as recorded in the tables, and as may be demonstrated through a Harris matrix, is a directed graph which may form nets, and not hierarchies, only.

Layers in the sense used here are two-dimensional views of three-dimensional phenomena. We tend to think of layers as three-dimensional realities, but so far we have seen no proper and practical means to excavate and document them as such, although there has been attempts in this direction (Reilly & Richards 1988; Stancic 1989). It is obvious that the layers that appear in our documentation are abstractions, and should not in any way be confused with our normal three-dimensional conception of layers in the field. Still they are the observed reality that we carry with us from the excavation.

We may, however, pose the question, whether it would not be preferable to aim at a three dimensional representation of the layers? Ideally yes, but there are some prohibitive problems involved. The resources needed on site to acquire the necessary raw data for a true three dimensional model, and afterwards to edit and correct the data, will be sizeable. The excavation strategy used should preferably be a top down layer by layer removal, which is a technique that cannot be used on prehistoric sites on glacial soils. The differences here between layers are often so subtle that they can be defined in sections only, and thus we are faced with the problem mentioned above, where we never see the layers as three-dimensional formations, but only as horizontal and especially vertical, two dimensional representations. Anyway, it is difficult to imagine that Paul Reilly's vision (Reilly 1991) of a solid model solution for excavation recording ("Virtual archaeology") can make an information management system like the one discussed in this paper superfluous.

Movable objects

In a strict logical sense it is not possible to separate between moveable objects and layers, as the latter fully incorporate moveable objects, and ultimately consist of moveable objects. Objects in the widest sense can be viewed as the basic building blocks that form the layers (Stein 1987: 339). To acknowledge this we need only remember that walls, posts and large stones in constructions, something that most excavators would prefer to view as layers, can be considered to be moveable objects in their own right, and indeed were so at one time. Likewise, any sample - even a 100% sample - of a layer automatically becomes a moveable object the minute we put it in a bag, and bring it with us. A moveable object is thus what we define to be a moveable object: that is for all practical purposes, anything - artifact, ecofact or sample - that we stick a label with a number on and bring with us as a whole.

Each and every object naturally belongs to a layer somehow, and indeed no registration should be allowed, unless a reference to a layer is also given. Ideally this ought to be a straightforward 1:M relation, meaning that any object should be located in one and only one layer, whereas of course a layer may contain no, one or many objects. However, it is not quite that simple. It may often be impossible to decide in practice whether an object belongs to the one or the other layer, i.e. situated at the interface or in cases of uncertain recovery. Further an object may crosscut the boundary of several layers either because it stood upright as the layers formed around it, or because it was stuck into the ground crosscutting layers already existing. As a consequence there is a N:M relation between objects and layers.

We may look at the moveable objects in their own right, and archaeologists certainly do so with artifacts. This, however, should not make us forget that moveable objects are extracts of layers, and that post excavation studies of objects are bound to add new information concerning the layers. This new information may either relate directly to the observed physical properties in the layer record, as is often the case with samples taken from the layers, or they may relate to the inferred information in the construct record, where it may give rise to alterations. As a rule the study of moveable objects will add to the knowledge of synchronism between layers, between constructs, and between layers and constructs, as well as knowledge concerning the function of constructs. A plotting of objects on sections or plans may result in an inferred splitting or aggregation of layers (Stein 1987: 343).

Added information obtained from moveable objects is also provided through a study of relations between objects. This, however, is sadly overlooked in most registration schemes for archaeological excavations, where moveable objects in most cases are considered as isolated units of information with some provenance information relating to the contextual record. The post excavation analyses, however, add an entirely different set of relations to the record. Thus when the artifacts have been cleaned, a process of refitting normally takes place. Fragments of what used to be complete utensils are fitted and glued together - notably pottery, but also flint waste and broken metal artifacts. This refitting of artifacts results in direct physical links between objects so that any object may relate to one or more other objects in different ways.

The relations between the objects are of three different types. Given a number of sherds that can be glued together to a part of a vessel, we may wish to know which sherds "tie up" with which, in order to reconstruct the pattern of destruction, but we may not wish to nominate the refitted sherds as an object. On the other hand, if we have two fragments of a flint axe that fit together to form one complete axe, we may of course wish to note that the two fragments "tie up with" each other, but we may also wish to register the axe as an object, and state that the two fragments "belongs to" this axe. Finally, if we glue together a flint flake material in order to reconstruct the production sequence, we certainly will wish to know, in addition to the above, which flakes are "above" which in the sequence. There is a very clear resemblance between the latter recording, and the stratigraphical recording held in the layers with layers table. To realize this you need only look at an example of a drawing of a refit given by Cziesla (1990:112). Its structural resemblance to a stratigraphical drawing is striking.

It follows that the range of possible relationships between moveable objects is of a more complex character than those between layers. In graph-theoretical terms, the relationships between objects can in some cases constitute "coloured" or "valued" graphs, i.e. different types of relationships are possible between individual pairs of moveable objects, whereas only one type of relationship can exist between any two layers. But apart from this difference it is possible in both cases to distinguish between directed and non-directed relationships. Thus the relationship "above" used in connection with both layers and moveable objects possess an implied direction, which has its inverse ("below") and is of a transitive nature (if A is "above" B, and B is "above" C then it follows that A is "above" C) (Orton 1980:66-80). Another example of a directed relationship between objects is "belong to" (sherd A belong to pot B) with its inverse "consist of". An example of a non-directed relationship is "tie up with" which by nature is symmetric (if sherd A "ties up with" sherd B, then it follows that sherd B "ties up with" sherd A.).

As stated above, the definition of what constitutes a moveable object is quite simple. It is anything we stick a number label to and bring home with us. As a consequence, moveable objects may be of many different kinds that call for different treatment. Therefore, an important part of the object record is the possibility of dividing the objects into categories and giving differing descriptions to these.

On a subsequent level of the post excavation analyses we may find a variety of information that links objects together in a non physical way. There is information that links with almost the same strength as the physical links mentioned above, and there is information that produces weaker links. Individual pots may be separated by comparisons of ware, temper and decoration, and groups of artifacts may also be separated using common technical elements. For instance, pots can be so identical that the same potter may be assumed, and technical investigations can show the same punch to have been used to decorate several metal artifacts. On a lower level, we may also categorise the objects with respect to some common elements that we find important in relation to the understanding of the excavation, whether they are technique, form or decoration elements.

These relations obviously cannot be represented internally in the object entity, as the groupings we create are not themselves objects. In the same way as we needed constructs to handle the inferences about layers we need constructs to handle groupings of objects based on various criteria. It could perhaps be argued that constructs used in connection with artefacts should be kept separately from constructs used in connection with layers. However, as we have found no logical differences in the way constructs are used with layers and with objects, and as we have not been able to find any drawbacks in keeping the information together, we have opted for a common construct recording. The construct dealt with in the next section thus covers layers as well as objects.

Adding meaning to layers and objects - the construct concept

Constructs can be considered as statements of abstract properties ascribed to layers and objects. That is, apart from being an entity that relate layers and objects in various combinations they also, by definition, add an inferential quality to the relation by way of a value label. You may think of the construct as a multi dimensional cross reference table with a label field, where you can add a value to the cross reference. Any construct may refer to layers, objects and other constructs in any combination and assemble these to a unity with some implied meaning.

It should be noted that the construct entity does not possess any inherited structuring categories. Concepts like interfaces, features and structures, so prominent in many systems, are to be considered interpretations only, and as such they have to be introduced deliberately as constructs. The construct entity is an open ended, neutral structuring facility, where the structure entirely depends on the archaeologist using the system. He may choose to use the constructs to simulate Harris's system, he may use it to simulate the *lag/anlæg* system used in Denmark, or he may use it as it is intended - as an open system where structures are created dynamically and fitted to the site being excavated.

Naturally, the central role of the construct is to give meaning to the layer recording. This is the major aim of any excavation. As interesting as the physical structure of the site may be, the interpretation assigned to the structure is what we are looking for in the end. The Harris matrix, eminently representing the physical structure, is thus not the end of analysis; it has to be interpreted properly. One could argue, as has been customary in excavation registration systems so far that the interpretation could be assigned to each layer as an attribute i.e. "Layer X represents a posthole". This, however, seriously limits our possibilities of handling the multidimensional nature of our data and interpretations. Hence, we argue that it is more appropriate to treat the interpretation as a separate entity.

In relational terms we need two tables to represent any combination of interpretations assigned. They are a table for the constructs proper, and a table for the relations between constructs. This construction overcomes the limitations mentioned in the introduction of fixed levels in hierarchical structures, and at the same time opens up for multiple interpretations of the same basic field-recordings in the simplest possible way.

The construct concept, however, is not only applicable in connection with the interpretation of the layers. It can also play a significant role in connection with objects in the post excavation process. A major part of writing up an excavation is the analytical work on the artifact material. Artifacts are categorised in different ways, and lists and maps are produced showing the distribution of artifacts belonging to various categories. This work is normally done outside the excavation recording system, working on extracts from this, but there is really no reason why this should be so. The categorisations may as well be an integrated part of the recording system.

The situation is very like the one encountered with layers. The categories may of course be added as attributes to the object recordings, but this would be a very unsatisfactory solution, as the categorisations of artifacts are truly multidimensional. That is, categories will cross cut each other in the artifact material in many different combinations, and every object may belong to several different categories. Here again, in relational terms we need two tables to represent any combination of interpretations assigned. We have chosen to use the same two tables for the object constructs as used for the layer constructs, because there are neither theoretical nor practical differences in the way the constructs relate to the two different entities.

It may at first glance seem a little odd that we should combine the analytical (inferential) analysis of layers and objects in this way. We are convinced, however, that on further thought it will be generally acknowledged as the only reasonable thing to do. And it does emphasize one of our major points: that we must keep interpretation strictly separated from observation. By using constructs to categorise object information, and to have these constructs fully integrated with constructs related to layers, we are provided with a very strong tool for the interpretation of an excavation. As an internal dynamic function to the excavation record, integrated with the context information, we can perform studies of synchronic and diachronic relations, and through these add new information to the record.

One of the more powerful features of the construct entity, which could not be, if interpretations were mere attributes to layers and objects, is the ability to operate with constructs of constructs. In complex sites this can become a very powerful inferential tool. Take the Vorbasse Iron Age village site for instance (Hvass 1988), where many thousands of postholes and foundation trenches are combined into many hundreds of houses and fences, further combined into hundreds of farmsteads, finally combined into tens of village phases covering more than a millenia. The work is currently done by hand, which means that every time a proposition for a house, a farmstead or a village is made, the records has to be meticulously checked for violations of the stratigraphical record, and the amount of checking grows rapidly with the level of integration. In a system, where constructs are implemented as suggested here, it will be possible for the archaeologist tentatively to suggest interpretations, and automatically have these checked for violations of the primary observations.

One final feature of the construct entity should be emphasized. This is the possibility of creating more holistic inferential units by combining layers, objects and constructs into one new construct. Take for instance an Iron Age house. This would as suggested above be defined as consisting of a number of other constructs, notably postholes and foundation trenches. These constructs would in turn individually be defined as consisting of a number of layers. However, there may be more to a house than the construction details. The house may contain artifacts that we would like to attribute to it, and there may be separate layers (a layer of ash around a fireplace for instance) that we would like to attribute as well.

The properties of excavation documentation entities

The layers, objects and construct entities discussed in the previous chapter comprises the basic archaeological informations relating to an excavation. And although it is only three entities (please note that we have not discussed the descriptive attributes needed in connection with layers and objects), the total design is not at all simple as we shall see below. However, there is more information to an excavation than these basic entities. There is also a complex set of entities used to document the basic information. Thus three dimensional measurements, drawings, photos and text notes are used as standard documentation in any modern excavation.

The need for text notes, we will argue, will disappear in an information management system as the one discussed here. Their function as tools to create structure and inference will be taken over by the facilities of the system, and whatever descriptions they may contain should find their way into the proper descriptive fields of the system.

In most excavation recording systems, measurements are viewed as attributes to objects. Any object can have one, and in some systems more than one set of measurements (x,y,z) attributed to it. The question of three dimensional measurements, however, is not all that simple. Any object found in an excavation, and indeed any layer defined, constitutes a three dimensional form in space. However, our recording on plans and sections, and any number of measurements that we attribute to an object or a layer (recorded objects attached to a layer may here add some information concerning the three-dimensional form of the layer) are only an approximation to its form and position in space. Thus ideally, any object and layer should be pin-pointed in space through an array of points sufficiently large to give an approximation to the form and position of the object or layer, but in practice, this ideal is difficult to accomplish. Photos constitute in this respect a further complication, because they are not intended to be exact positioned recordings, but rather image recordings of the excavation in progress.

Despite the difficulties, the idea of an implementation and presentation of a three dimensional model of an excavated site should not be discarded, but in the context of this article we have refrained from attempting a proper logical design of a measurement entity. This is not as much because of the problems with missing and bad data, and problems with transfer of drawings and photos into structured (x,y,z) information, as it is because the relational database model, which has been fully adopted in this paper, does not fit the nature of three dimensional modeling. We are of course aware that excavation recording systems with capabilities of three-dimensional form recording exist, but in order to fulfil the very important demand of independence between the database proper and the generation of applications to the database, we will not go on compromise with the design. A possible solution to the problem, to be investigated as a next step, may be found with object oriented database design (Williamson & Stucky 1991).

As a consequence of these shortcomings, we have decided for the time being to view measurements the traditional way - as attributes to layers, objects and constructs. We thus deal with two separate entities for excavation documentation only: drawings and photos. It follows that the relations between these entities has to be viewed as relations between uniquely named or numbered layers, objects, drawings and photos, where the names or numbers serve as reference identifiers for the entities and serve at the same time as navigation tools between them.

Drawings

Drawings of plans and sections are no doubt the most important of our means of documentation. They are used to document the position and form of both layers and objects. Mostly, they are used in connection with layers, and indeed it is customary that all layers recorded on an excavation appear on at least one plan or section drawing. Objects need not, and in many excavations never appear on the drawings. In some excavation contexts, however, it is customary to show the objects on plan and section drawings, recognizable as individual objects and with clear information concerning their deposition (note we are not speaking of point plots of object positions). In these cases the objects have exactly the same relations with drawings as do layers.

As one layer or one object may appear on many drawings, and as one drawing may reference many layers and many objects, we have true N:M relations between drawings and layers and objects respectively. Further, as it is necessary to cross-reference drawings in excavations, and indeed as it is customary to indicate the position of other drawings on drawings, it is necessary to have a table that handles drawings with drawings. Finally, as photos almost always are cross referenced with drawings it is necessary to have a table that relates drawings with photos.

Photos

Photos are used intensively to document layers as well as objects. Many archaeologists are quite sloppy about cross referencing their photos properly with layers and objects, but that does not mean that there is no reason to consider this problem. Just as with drawings there is a true N:M relation with both layers and objects. Any layer or object may be found on many photos, and any photo may show many layers or objects. Further there is the relation with the drawings mentioned above.

A basic design model for an excavation registration system

The design model

Our design model is presented in Figs. 1 and 2. Unfortunately, the use of two figures is necessary to avoid a web of crossing lines. The first figure (Fig. 1) shows the logical design for the three basic entities: layers, moveable objects and constructs with their internal and external relationships (for the concept of internal and external relationships see le Maitre 1981). The various levels of the model are indicated through different shadings. The pointed arrows indicates that each record in a table may point to N other records in the connected table.

The second figure (Fig. 2) shows how the documentation entities, drawings and photos, relate to the basic entities, layers and moveable objects, with their internal and external relationships as well. Due to the presence of four entities a single crossing line is unavoidable, but the diagram should be readable all the same.

Exemplification of the model

In order to make our model more readily understandable we shall present two examples. The first example is based on information from the publication of the Hodde Iron Age village

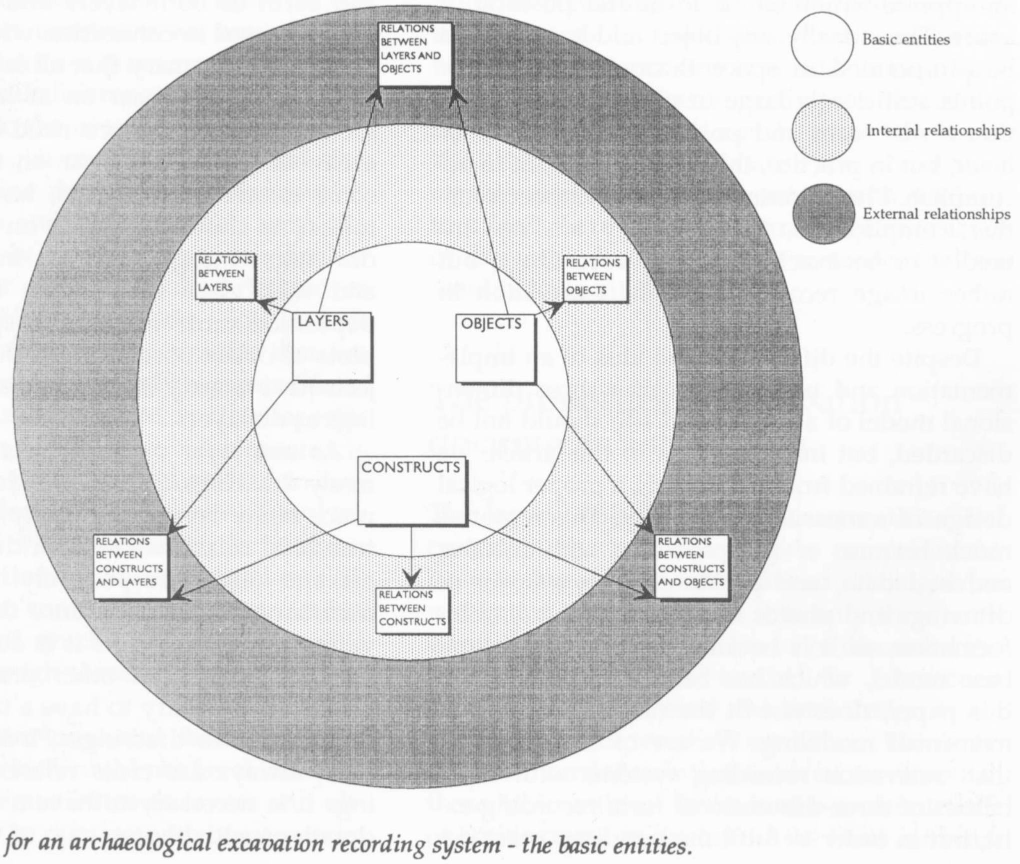

Fig. 1. Design model for an archaeological excavation recording system - the basic entities.

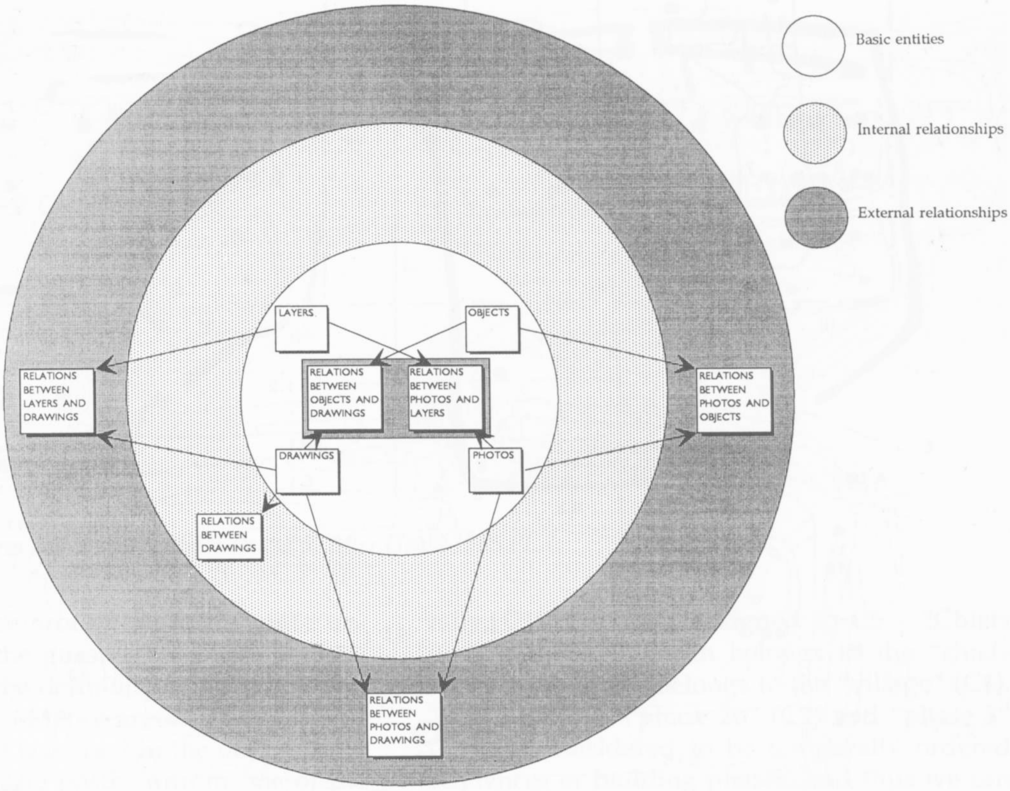

Fig. 2. Design model for an archaeological excavation system - the documentation entities.

(Hvass 1985). It shows how complex contextual information can be handled through the use of the construct concept. The other example is based on informations from the publication of a palaeolithic settlement at Gönnersdorf in the Middle Rhein area (Bosinski 1979). This example shows how complex relations between objects in an excavation can be handled.

The Hodde site is a typical older Iron Age village from Jutland. This goes for the layout of the village as well as the way the site presented itself during excavation. The village is surrounded by a fence along which the individual farms are placed. Some of the farms are also surrounded by fences. The village had been situated on the same spot for a period of time sufficiently long for most of the houses to have been renewed and the fences reset. Archaeologically, all information concerning the village, and its various reconstruction phases is only preserved as a crisscross maze of cuttings in the subsoil - cuttings that primarily stem from fences, post-holes and various types of pits. To understand the village, it is necessary to build a huge inferential model that states what goes with what, and what is contemporary with what.

In our conceptual scheme such a model is to be found entirely at the level of the construct, and only at its basic level does it rest on layer information. This layer information is numerous, but typically very fragmented. That is to say, the sequences of layers that relate stratigraphically to each other are very short but they occur in abundance.

As a typical example we may take the so-called "chiefs farm" at Hodde (Fig. 3). We see several fences and buildings superimposed - too many in fact for us to make a complete listing of the tables here. We will confine ourselves to a few listings to show how it would work.

As an example of the type of stratigraphic information that prevails at Hodde, Fig. 3 shows a plan of the "chiefs farm" with *anlæg* numbers written on, and a couple of sections through a set of fences and a posthole with *lag* numbers written on. The section through A,14 is in fact a section through two overlaying fences of which

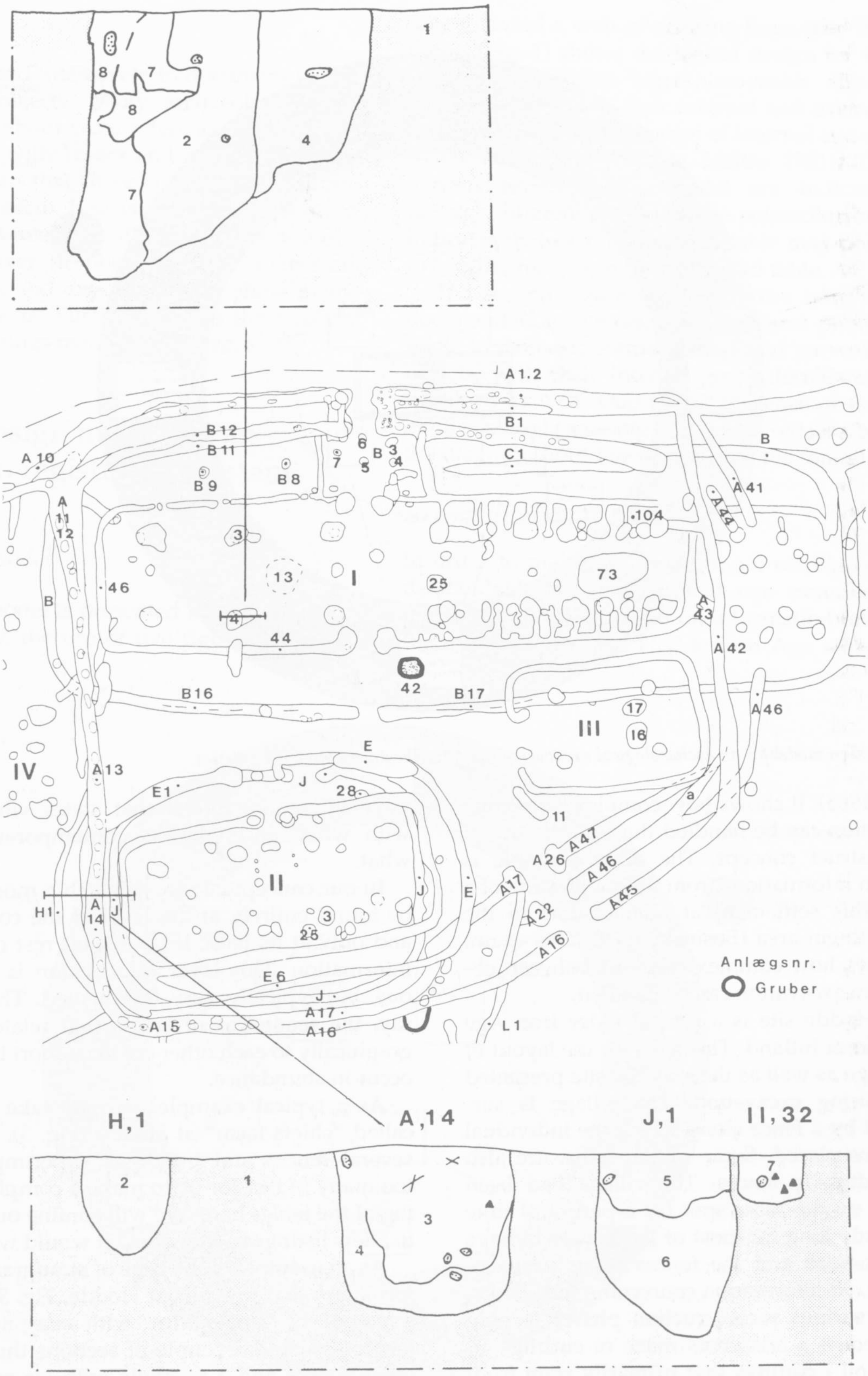

Fig. 3. Plan of "chiefs farm" with anlægs *numbers shown and sections through a posthole and a set ofences with* lag *numbers shown. After Hvass 1985 Pl. 4, Fig. 8 and Fig. 20.*

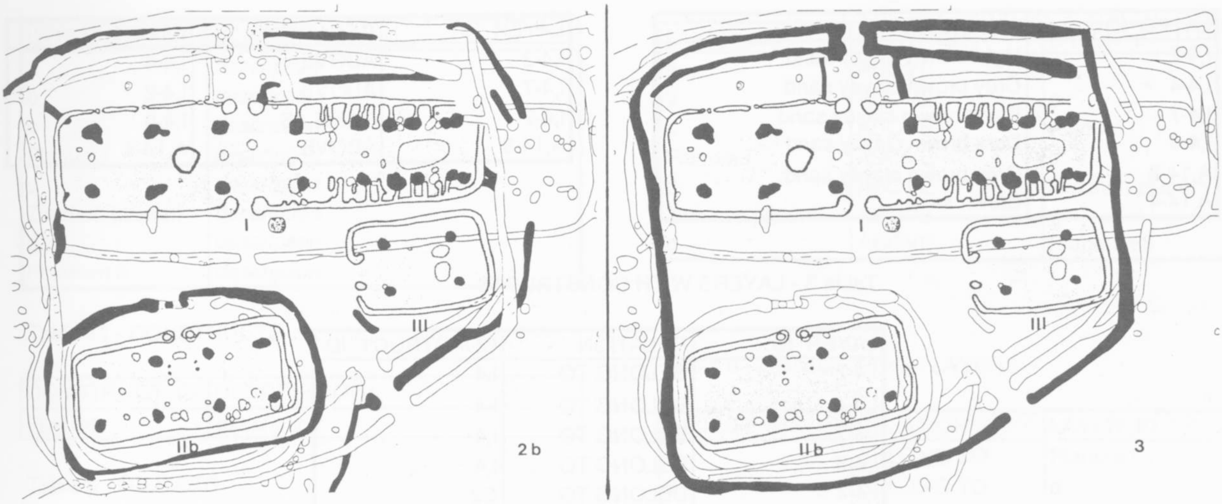

Fig. 4. The "chiefs farm" at Hodde in phase 2b and 3. After Hvass 1985 Pl. 7.

the lower (a fence around what we quite unwarranted term the guest house) was used by the excavator in the definition of phase 2b (Fig. 4), and the upper (a fence around the whole of the "chiefs farm") was used in the definition of phase 3 (Fig. 4). The posthole from one of the roof-carrying posts of the main building has been reset once, and we claim (the excavator may indeed disagree with us) that the earlier post stems from phase 2b, while the later is from phase 3.

The two sections on Fig. 3 gives us the tables shown of layers (Table 1), layers with layers (Table 2) and layers with constructs (Table 3). In the latter table we may note with reference to Table 4 that I,4 is a construct named "posthole", C2 is a construct named "phase 2b", C3 is a construct named "phase 3", A,14 and E,1 are constructs named "Fence section". Indeed, the constructs table (Table 4) show how we can name almost anything on very different levels of integration. We have very basic stratigraphical constructs like "fence section" and very high level constructs like "village".

The Constructs with Constructs table (Table 5) shows how this information subsequently can be combined and integrated. Part of the "fence section" is grouped into C8, which is "Fence around chiefs farm", the rest into C9, which is "Fence around chiefs guest house". Further, we find that I,3 and I,4, the two postholes, belong to a construct, C10, named "roof carrying posts in main building". Naturally, C10 is further assigned to C5 - "Chiefs main building", which belongs to the "chiefs farm" (C4), which belongs to the "village" (C1). The Constructs "phase 2b" (C2) and "phase 3" (C3) are considered to be temporally ordered sequences of building phases, and thus we can state that C3 is above C2. Further we may assign various constructs to the two phases. "Fence around guest house" (C9), is thus assigned to C2, and "fence around chiefs farm" (C8) is assigned to C3.

The Gönnersdorf site is basically a Magdalenien living floor including the remains of a hut sandwiched into a deeply stratified loess sediment. The loess sediment has protected the living floor, leaving this as it was when left by its occupants. On fossil living floors like this, much information can be gained by studying how the objects relate to each other on the floor.

One of the more interesting aspects of palaeolithic or mesolithic sites is the production and consumption patterns of artifacts as revealed specifically through refitting of the leftover material on the site. It may give information on how artifacts were produced, where they were produced, what kind of artifacts were produced on the site, and even what artifacts were removed from the site and deposited elsewhere.

We have chosen to show a minor refitting, and outline what this means in terms of relations to be registered. It should be noted that

Table 1 - LAYERS

LAYER ID	DESCRIPTION
I,4-2	Light brown clayey sand
I,4-4	Grey brown clayey sand
I,4-7	Black brown clayey sand
I,4-8	Dark brown clayey sand
A,14-3	Grey brown clayey sand
A,14-4	Brown sand

Table 2 - LAYERS WITH LAYERS

LAYER ID	RELATION	LAYER ID
I,4-2	ABOVE	I,4-4
I,4-7	ABOVE	I,4-2
I,4-7	SAME AS	I,4-8
A,14-3	ABOVE	A,14-4

Table 3 - LAYERS WITH CONSTRUCTS

LAYER ID	RELATION	CONSTRUCT ID
I,4-2	BELONG TO	I,4
I,4-4	BELONG TO	I,4
I,4-7	BELONG TO	I,4
I,4-8	BELONG TO	I,4
I,4-4	BELONG TO	C2
I,4-2	BELONG TO	C3
I,4-7	BELONG TO	C3
I,4-8	BELONG TO	C3
A,14-3	BELONG TO	C3
A,14-4	BELONG TO	C2
A,14-3	BELONG TO	A,14
A,14-4	BELONG TO	E,1

Table 4 - CONSTRUCTS

CONSTRUCT ID	DESCRIPTION
C1	Village
C2	Phase 2b
C3	Phase 3
C4	Chiefs Farm
C5	Chiefs main building (I)
C6	Chiefs storage house (III)
C7	Chiefs guest house (II)
C8	Fence around chiefs farm
C9	Fence around guest house
C10	Roof carrying posts in main building
C11	House wall
I,3	Posthole
I,4	Posthole
A,11	Fence section
A,12	Fence section
A,13	Fence section
A,14	Fence section
A,15	Fence section
A,16	Fence section
B,12	Fence section
A,1.2	Fence section
A,42	Fence section
A,45	Fence section
E	Fence section
E,1	Fence section
E,6	Fence section

Table 5 - CONSTRUCTS WITH CONSTRUCTS

CONSTRUCT ID	RELATION	CONSTRUCT ID
C3	ABOVE	C2
C4	BELONG TO	C1
C5	BELONG TO	C4
C6	BELONG TO	C4
C7	BELONG TO	C4
C8	BELONG TO	C4
C8	BELONG TO	C3
C9	BELONG TO	C7
C9	BELONG TO	C2
C10	BELONG TO	C5
C11	BELONG TO	C5
I,3	BELONG TO	C10
I,4	BELONG TO	C10
A,11	BELONG TO	C8
A,12	BELONG TO	C8
A,13	BELONG TO	C8
A,14	BELONG TO	C8
A,15	BELONG TO	C8
A,16	BELONG TO	C8
B,12	BELONG TO	C8
A,1.2	BELONG TO	C8
A,42	BELONG TO	C8
A,45	BELONG TO	C8
E	BELONG TO	C9
E,1	BELONG TO	C9
E,6	BELONG TO	C9

Table 6 - LAYERS

LAYER_ID	DESCRIPTION
a	*description*
b	*description*
c	*description*
Planum I	*description*
d	*description*
Zwichen planum	*description*
d	*description*
Plannum II	*description*

Table 11 - CONSTRUCTS

CONSTRUCT_ID	DESCRIPTION
Flint refit	*description*

Table 9 - OBJECTS

OBJECT_ID	DESCRIPTION
7	Core
9	Flake
48	Flake
76	Flake
78	Flake
88	Flake
999	Core

Table 7 - LAYERS WITH LAYERS

LAYER_ID	RELATION	LAYER_ID
a	ABOVE	b
b	ABOVE	c
c	ABOVE	Planum I
Planum I	ABOVE	d
d	ABOVE	Zwichen planum
Zwichen planum	ABOVE	d
d	ABOVE	Plannum II

Table 8 - OBJECTS WITH LAYERS

OBJECT_ID	RELATION	LAYER_ID
7	BELONG TO	Planum I
9	BELONG TO	d
48	BELONG TO	d
76	BELONG TO	c
78	BELONG TO	d
88	BELONG TO	c

Table 12 - OBJECTS WITH CONSTRUCTS

OBJECT_ID	RELATION	CONSTRUCT_ID
999	BELONG TO	Flint refit

Table 10 - OBJECTS WITH OBJECTS

OBJECT_ID	RELATION	LAYER_ID
7	ABOVE	88
9	ABOVE	78
48	ABOVE	9
76	ABOVE	48
7	BELONG TO	999
9	BELONG TO	999
48	BELONG TO	999
76	BELONG TO	999
78	BELONG TO	999
88	BELONG TO	999
7	TIE UP WITH	48
7	TIE UP WITH	76
7	TIE UP WITH	76

refitting of flint not only give informations as to which pieces tie up with each other, but also in many cases information on the sequence of detachment in the production process. There is so to speak a stratigraphy among the refitted pieces.

The Gönnersdorf site was excavated in a number of partially natural and partially artificial strata to which the artifacts were assigned. Three phases of the living floor were defined as well, and artifacts were assigned to them. The stratigraphy, as shown in a greatly simplified section Fig. 5, results in the two Tables 6 and 7 of Layers and Layers with Layers.

The refit we have chosen (Fig. 6) consist of six pieces of quartz, each of which was assigned to a layer, as shown in Table 8. It should be noted that the numbering of the objects used here is our own shorthand version of the original rather complex numbering system used at the excavation.

The position of the objects on the floor is shown on the plan Fig. 7. They amount to one core and five flakes. Lines indicate how the pieces fit together. Lines with arrows mean an ordered "above" sequence, whereas lines without arrows indicate that the pieces just "tie up with" each other.

In the Objects (Table 9) and Objects with Objects (Table 10) tables we can see how this information may be handled. Note that we have added a core 999, which is the core that comes out of the refit. As a consequence all six objects belong to this core. In the final two tables (Tables 11 and 12) we demonstrate how we may define a construct Flint Refit of which core 999 will be a member.

Conclusion

It is not generally realized, we think, that the logic used to organize information is to a large extent determined by the tools available to us for organizing and storing the information. We simply tend to think along the lines that our tools allow us. Using pen and paper this simply implies flat file organization.

Our view of the anatomy of an archaeological excavation, as we have presented it here, is strongly influenced, of course, by relational information system concepts, simply because the tool we intend to use is a relational database system. There may be flaws in our scheme, and better schemes may turn up. Nor are we certain that this is our final bid. What is final, however, is the shift from flat file to relational file systems, or what may turn up beyond in terms of object oriented database systems etc. It is a much more powerful way to organize and handle information than is the former. Given powerful computers and relational database systems it is a waste of effort to continue with flat file systems.

As a consequence, we would urge those who are responsible for the registration of excavation data to reconsider the structure of their data, and to do this from the point of view of relational information structures. We are convinced that it is a matter of a few years only, before the existing registration systems based on the flat file concept will have either to be discarded entirely or restructured fundamentally.

At the point of this major alteration in the registration systems, it may be worth while to consider, whether we should accept or perhaps even encourage the current state of diversity among systems to continue with the new systems, or whether we should actively try to standardize the use of recording systems so

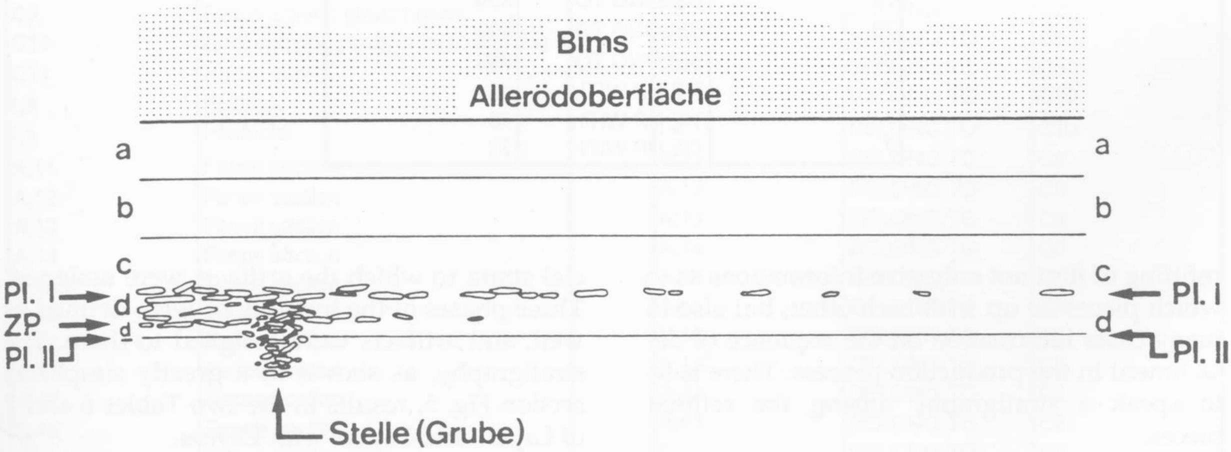

Fig. 5. The stratigraphy at Gönnersdorf. After Bosinski 1979 Abb. 19.

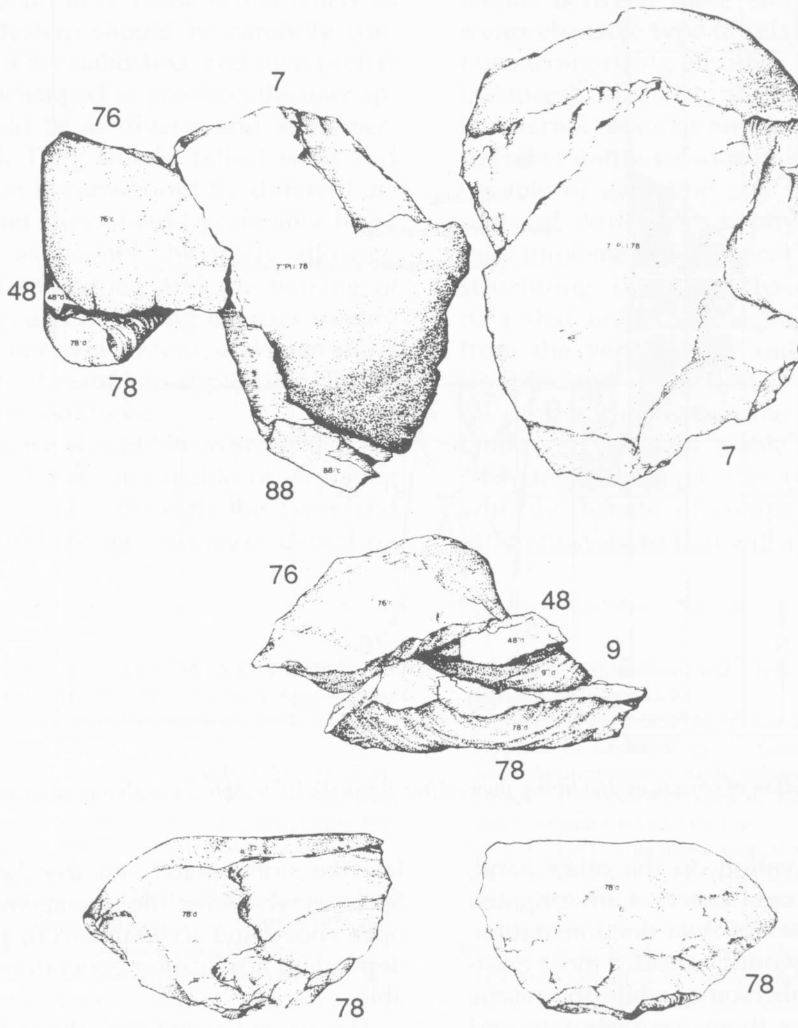

Fig. 6. The refitted object. After Bosinski 1979 Abb. 53 with our addition of numbers.

that only one or a very few are in use? The latter position has recently been taken by Arroyo-Bishop (1989, 1991), but although we accept the benefits of standardization we feel it depends very much on what is standardized. The question, really, should be split in two, one being: should we standardize the design of the logical structure? and the other: should we standardize the user interface and its interactions with the database? The answer to the first question is yes, the answer to the second is no.

As for the first part of the question, the logical design should reflect our theoretical considerations of the structure of archaeological data, and to the extend that we agree on these, we should also be able to agree on the logical design. It is obvious that we may experience disagreement, and we may end up with different "schools of thought", but on the whole we will probably see much more conformity than so far experienced if we involve ourselves in an explicit discussion.

The second part of the question concerns the user applications that carry out the transactions with the file system based on the logical design. These should reflect local, individual needs closely associated with the excavation framework within which the archaeologist is working. Take as an example two very different excavation situations. The research excavation on the one hand, where elaborate documentation means more than volumes of earth removed,

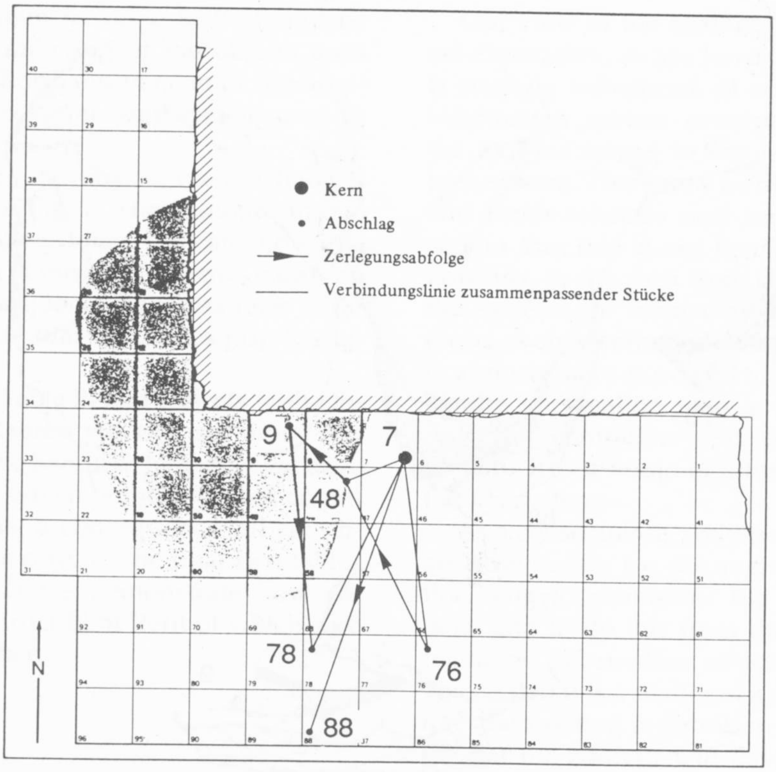

Fig. 7. Plan of the association of objects on the living floor. After Bosinski 1979 Abb. 54 with our addition of numbers.

and the rescue excavation on the other hand, where price per square meter of investigated surface is more important than documentation. The research teams would consider most existing systems too cumbersome, while the rescue team would consider them too elaborate and expensive to use. This situation would not change just because new systems are designed. Yet the sites the two teams are digging may be very much alike - it could even be the same site - so it seems logical that the data they register should be held in the same database. Obviously, two very different systems are called for, but only as far as the user interface and the interaction with the database is concerned. What is needed is a basic database structure that can hold and organize even very elaborate documentation, and yet is so simple and flexible that it can be used in rescue excavations, where documentation is scanty at best. We feel that the design presented here fulfil these needs.

Given the above considerations we would like to state a few points concerning the implementation of the system. The first we shall make is that there should be a separation between the applications that access and manipulate the stored data, and the data themselves. Systems where the file management part is not open ended and accessible from a number of independent application generators is not acceptable.

The second point to make is that great care should be taken in the design of the file management system. It must be as straightforward as possible, but it cannot be simple because the data to be handled are neither simple nor one-dimensional. On the contrary they are complex and multidimensional.

The third point follows from the second. The file management system used must be a very advanced one. It has to be truly relational, supporting not only 1:M and N:M relations between entities, but also within entities, and it has to support different types of relations. Such systems exist on the main frame level, but only partially in the low end of the computer market, and those that meet the demands are liable to die of indigestion, when they are put to the test in a really complex database application. But give it a few years. Powerful workstations, by then extremely cheap, will do the trick.

The fourth point to be made is that where as the basic file system should be carefully considered before it is established, and then preferably kept as unchanged as possible, the user applications should be as diverse and as numerous as needed. They should reflect what and how registration is carried out by different archaeologists, and they should preferably be as user tailorable as possible by freely allowing the creation of categories, and the naming of these on the fly, and by letting the user modify or create his or her own screens, and even allow for the creation of complete applications using modern development tools.

In conclusion we would like to stress that the design suggested here is capable of depicting any physical structure through the layer and object entities and the internal and external relations between these entities using the most comprehensive type of relations - the M:N relations (comprising all other types of relations as instances). Likewise, the construct entity with its internal relations and its external relations to the other entities also using the M:N relations is capable of depicting any interpretation of the physical structure. A physical design which fully implement this logical design can hold any structuring of the archaeological excavation data that an archaeologist may decide upon, from the very simple and crude to the very complex and sophisticated recording. The logical design implies that the physical implementation can be kept stable and untouched by changing demands for recording standards with the benefit of compatibility between the different systems that will appear.

References

Arroyo-Bishop, D., 1989: The ArchéoDATA Project. In Rahtz, Sebastian & Julian Richards (eds) Computer Applications and Quantitative Methods in Archaeology 1989. BA International Series 548 pp. 69-86.

1991: The ArchéoDATA System - towards a European archaeological document. In Lockyear, Kris & Sebastian Rahtz (eds) Computer Applications and Quantitative Methods in Archaeology 1990. BAR International Series 565. Tempus Reparatum. pp. 61-69.

Bosinski, G., 1979: Der Magdalénien-Fundplatz Gönnersdorf. Band 3. Die Ausgrabungen in Gönnersdorf 1968-1976 und die Siedlungsbefunde der Grabung 1968. Wiesbaden.

Carver, M.O.H., 1985: The friendly user. In Cooper M. & J. Richards (eds) Current issues in archaeological computing. BAR International Series 271. pp. 47-62.

Cziesla, E., 1990: Siedlungsdynamik auf steinzeitlichen Fundplätzen: methodische Aspekte zur Analyse latenter Strukturen. Studies in modern archaeology Vol.2. Bonn.

Harris, E.C., 1979: Principles of Archaeological Stratigraphy. Academic Press, London.

Hinchliffe, J. & Jefferies, J., 1985: Ten years of data-processing in the central excavation unit. In Cooper, M. & J. Richards (eds) Current issues in archaeological computing BAR International Series 271. pp. 17-22.

Hvass, S., 1985: Hodde. Et vestjysk landsbysamfund fra ældre jernalder. Arkæologiske Studier Vol. VII. København.

1988: Jernalderens bebyggelse. In Mortensen, Peder & Birgit M. Rasmussen. Fra stamme til stat i Danmark I. Jernalderens stammesamfund. Jysk Arkæologiske Selskabs Skrifter XXII, pp. 53-92.

Maitre, J. Le, 1981: SOFIA: A Data Base Management System Applied to Archaeology. In Gaines, Sylvia W. (ed.) Data Bank Applications in Archaeology. pp. 123-129.

Orton, C., 1980: Mathematics in Archaeology. William Collins Sons and Co Ltd, London.

Reilly, P., 1991: Towards a virtual archaeology. In Lockyear, Kris & Sebastian Rahtz (eds) Computer Applications and Quantitative Methods in Archaeology 1990. BAR International Series 565. Tempus Reparatum. pp. 133-139.

Reilly, P. & Richards, J., 1988: New perspectives on Sutton Hoo; the potential of 3-D graphics. In Ruggles, C.L.N. & S.P.Q. Rahtz, Computer and Quantitative Methods in Archaeology 1987. BAR International Series 393. pp. 173- 186.

Ryan, N. & Smith, D., 1986: Databases, data models and archaeological theory. Science and Archaeology no. 28: 56-60.

Spence, C. (ed), 1990: Archaeological site manual. Department of Urban Archaeology. Museum of London.

Stancic, Z., 1989: Computervision: producing intra-site plans. Archaeological Computing Newsletter, 20: 1-20.

Stein, J. K., 1987: Deposits fo Archaeologists. Advances in Archaeological Method and Theory, Vol 11: 337-395.

Williamson, R. & Stucky, J., 1991: An object-oriented geographical information system. In Gupta, Rajiv & Ellis Horowitz,. Object-Oriented Databases with Applications to CASE, Networks, and VLSI CAD. pp. 296-312.