EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2018)
R. Sablatnig and M. Wimmer (Editors)

*Short Paper*

# Integrated spatio-temporal documentation and analysis of archaeological stratifications using the Harris Matrix

W. Neubauer[1], C. Traxler[2], A. Lenzhofer[1] and M. Kucera[1]

[1]LBI for Archaeological Prospection and Virtual Archaeology, Hohe Warte 38, 1190 Vienna, Austria
[2]VRVis Forschungs GmbH., Donau-City-Strasse 11, 1220 Vienna, Austria

**Abstract**

*The Harris Matrix (HM) is the fundamental diagrammatic representation of relative time for an archaeological site and the de facto standard for the representation of a stratigraphic sequence – the backbone for archaeological stratigraphy. It displays all uniquely identified units of stratification in a sequential diagram representing their relative temporal succession. The Harris Matrix Composer is a widely used application in the archaeological community to efficiently create and analyse HMs. However, it does not support explicit dating of HM units, which is an important information for post-excavation investigations of an archaeological site. In this paper we describe an integrated approach for a combination of stratigraphic and chronologic relations. The implicit, chronologic sequence given by the HM becomes explicit as scientists are enabled to define a hierarchical time model and assign units of the HM to temporal intervals or provide exact dating. The system maintains a consistent visual representation, which means that a correct stratigraphic layout is preserved while units are aligned to intervals of the time model. Evaluation of a real-world use case showed that this combined visualisation makes the scientific analysis and interpretation more efficient and reliable.*

**CCS Concepts**

•***Human-centered computing*** → ***Information visualization;*** •***Information systems*** → ***Temporal data;*** *Spatial-temporal systems;* •***Applied computing*** → ***Archaeology;***

## 1. Introduction

Any archaeological site or landscape can be understood as a physical volumetric body with limited spatial extent formed over a long period of time. Usually, natural, anthropogenic and anthropogenic-influenced physical or chemical processes contribute to the formation of a site, reflecting its environmental, historical and cultural settings. These volumes are organized by individual volumetric entities, the stratigraphic units, forming a unique stratification [Har89] incorporating the spatial and temporal aspects of the site corresponding to distinct events and time intervals. Every stratigraphic unit can be characterized by its geographical position and extent, its observed topological relations in relation to the other units and its specific temporal characteristics.

The Harris Matrix (HM), introduced in 1973 by the Bermudian archaeologists Edward C. Harris [Har89], is the de facto standard for the documentation of the topological relations of the archaeological stratification during the stratigraphic excavation process. The HM is not a matrix in mathematical terms but a sequential diagram representing the stratigraphic relations of all individual units of stratification. The initial definition and layout of an HM has been developed further by the first author in accordance with E.C. Harris and lead to a new convention implemented in software named Harris Matrix Composer [TN08].

Such a revised HM or stratigraphic sequence consists of two distinct type of nodes, rectangular symbols (□) representing material deposits documented and removed by the excavation process and circular symbols (○) representing immaterial feature surfaces (pits, ditches etc.), whereby edges represent the topological relations also known as the stratigraphic relations between them [TN08, Neu07]. In mathematical terms, a stratigraphic sequence or HM is a directed acyclic graph with different nodes for the two types of stratigraphic units, i.e. deposits and surfaces where edges define the stratigraphic relation “is above”. A directed edge running from unit A to B means that A lies stratigraphically above B or in other words A is in superposition to B. This implies that A is later than B in respect to their relative temporal succession. If A and B are not in relation, their temporal position is open and not directly defined by the stratigraphic sequence.

Hence the stratigraphic sequence can be seen as a relative calendar but the exact dates of units cannot be derived from it. Moreover, parallel strands do not necessarily mean that their units belong to the same time period, i.e. the horizontal alignment of units has no chronological meaning. This is a shortcoming as clear relations be-



DOI: 10.2312/gch.20181369

tween stratigraphic and chronologic properties would allow a more efficient and accurate analysis and a more robust interpretation. The temporal relations are determined by post-excavation find analysis and scientific dating methods typically resulting in time intervals with certain probabilities. Allen [All83] introduced a sound theoretical framework for temporal reasoning also known as Allen's interval algebra. It is very well suited for chronology in archaeology as it considers imprecisions, uncertainties and relative relations.

To be able to fully reconstruct the excavated volume and to derive the respective topological relations exemplifying the relative temporal succession, we have initiated the documentation of the surfaces of all excavated deposits in 3D, a process we have named Single Surface Planning, an important theoretical and practical extension of archaeological stratigraphic theory [Neu07]. The reconstruction of a site through time, i.e., the creation of the surfaces of each respective period, is achieved by the selection of subsets of the recorded single surfaces and by compiling the single surfaces from the bottom of the site upwards through time.

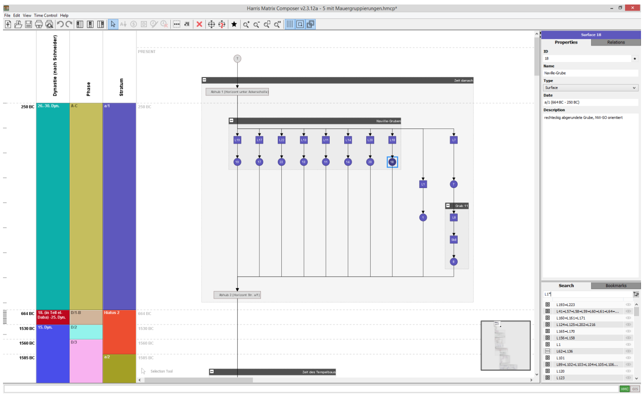

**Figure 1:** *Main user interface components of the HMC+*

## 2. Related Work

The work presented in this paper is based on the results of a previous research project, where a new editor for the compilation and validation of a stratigraphic sequence was developed, named Harris Matrix Composer (HMC) [HMC18]. The HMC is currently widely used in the archaeological community and even found application in forensics [IJ17, Han04]. The HMC is based on the powerful graph library yFiles from yWorks [yG18] and described in detail by Traxler and Neubauer [TN08]. The HMC can handle large directed acyclic graphs with a high degree of usability and efficiency.

Beside the stratigraphic or topological relations based on superposition, the HMC also supports temporal relations to define relationships of units of stratification that are not in superposition but can be defined in the way that A "is contemporary with" B or A "is later than" B. Additionally, it is possible to group subgraphs into structural entities called phases in the form A "is in phase with" B, C, .., X and into periods, assigning horizontal blocks of units to a historical epoch.

However, both of these methods proved to be unintuitive and were rarely used. Especially temporal edges are easily confused with stratigraphic ones impairing the analysis. For that reason, it was necessary to research on new concepts how to combine stratigraphic and temporal relations in one diagrammatic view.

## 3. The HMC+ – integrating an interval-based time model

### 3.1. Concept and design

To address the shortcoming of missing explicit chronological relations, we integrated a hierarchical time model into the HMC that enables dating of units. We investigated how stratigraphic and interval-based temporal relations can be combined in a consistent visual representation, resulting in the new tool named Harris Matrix Composer Plus (HMC+). The HM is therefore linked with the time model to assign units to time intervals. It is important that these time intervals can be fuzzy and overlapping as almost any archaeological dating method has uncertainties. Figure 1 shows the design of the user interface of the HMC+. The left panel contains the hierarchical interval-based time model, the middle one contains the graph editor for creating and analysing an HM and the right one other interface components such as the property editor or the search dialogue.

We implemented four concepts to represent the hierarchical time model that can be used to create respective temporal elements/objects derived from archaeological practice. Initially, there is an absolute time axis with a resolution set to one year. The initial set-up of the left panel provides three empty time frames with increasing levels of granularity: ages, periods and phases. Every individual unit can be directly assigned to a distinct time interval. Ages represent the top level of the hierarchy and depict coarse time frames that correspond to long historical epochs, such as the Neolithic, the Bronze Age and Iron Age. Periods are subdivisions of ages, such as the Early, Middle and Late Neolithic (LN) or even broken down to smaller entities. Phases depict the bottom level of the hierarchy. They are site-specific time intervals resulting from the analysis of the stratification and the dating of finds and samples.

All time frames and their respective subdivisions can be created and revised continuously by the archaeologist in the course of the post-excavation analysis. Per convention ages and periods do not allow for overlapping intervals. This means that every new element in the respective time frame, except the first one, must share a common border with an existing temporal element. In contrast, temporal elements in the phases time frame can overlap or have gaps between each other. This means that two consecutive elements must satisfy one of the following Allen relations: "x before y", "x meets y", "x overlaps y".

The HMC+ offers two different modes to date a unit: First, by assigning the individual unit to a defined interval of one of the three time frames or second by directly specifying an individual and distinct time interval, which is usually derived from archaeological or scientific dating methods such as $^{14}C$.

### 3.2. Layout Calculation

The visual representation of such a time model and the HM mutually affect each other. Units in the HM have to be vertically rearranged to align them with an assigned time interval. The stratigraphic relations must not be destroyed by this rearrangement. On the other hand, the time intervals need to grow visually to encompass one or more sections of the graph with equally dated units. Therefore, the time model is non-linearly distorted, i.e. the size of intervals does not necessarily correspond to their duration. Horizontal dotted lines at the upper and lower bound of the phase elements indicate the absolute time range.

Maintaining a linked view between the time model and the HM in real-time is not trivial as HMs are usually huge, and many cases have to be discerned. Allen's Interval Algebra [All83] was very helpful to discern different cases occurring when combining stratigraphy with a hierarchical time model.

The dating information provided is then utilised by a customised Sugiyama [STT81] layout algorithm to shift units vertically up or down the trail to reflect the corresponding time interval (depending on their assigned time element or date). However, the relative order of units in the sequence is preserved.

The approach can be decomposed into three phases: Layer Assignment, Crossing Minimization, and Coordinate Assignment [JM04]. In the first phase, each node in the graph is assigned to a layer, whereby two adjacent nodes may not be in the same layer. The minimum number of layers results from the length of the longest directed path in the graph.

To improve the readability of graphs, Sugiyama and Misue defined five prioritised rules for a structural drawing [SM91]. The first rule with the highest priority is called closeness and aims for a compact layout, which means that two connected vertices should be placed as close to each other as possible. So, to guarantee a minimum length between a node X which is assigned to layer i, and its successor Y, Y should ideally be assigned to layer i+1.

Length minimisation was also formulated as an aesthetic criterion for human readable graph drawings by Fleischer et al. [FH01]. However, this aesthetic criterion is broken in favour of a vertical rearrangement based on temporal properties, which can lead to longer edges between two adjacent nodes. Therefore, the first phase was modified considering Allen's Interval Algebra to achieve a partitioned, temporal ordered layout. Figure 2 shows how time intervals are stretched to enclose a vertical strand of dated stratigraphic units.

The following conventions are considered for the layout algorithm:

- Stratigraphy overrules chronology
- Every unit, except top and bottom surface, must have at least one predecessor and one successor
- A dated unit is always assigned to an interval, never to a specific point in time
- An undated unit is placed in the same time track as its oldest dated predecessor
- The layout is oriented from top to bottom. The arrow of time is pointing in the opposite direction
- Units with a custom interval are drawn with a thick border.

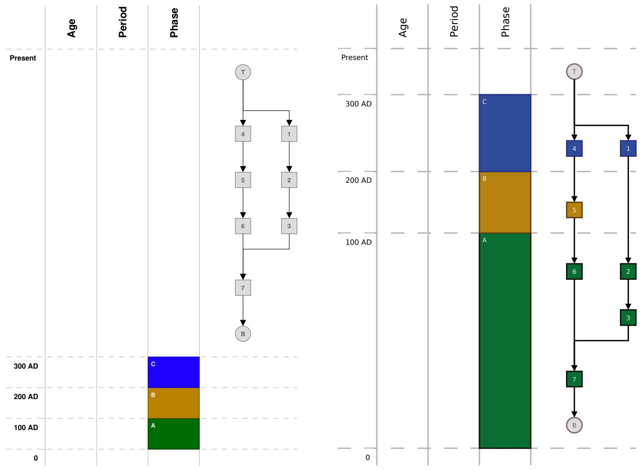


**Figure 2:** *Left: An undated HM with three phases, each lasting a hundred years. Right: The same HM, but dated. The phases are stretched to enclose their corresponding dated stratigraphic units.*

### 3.3. Colour Coding

In combination with a colour coding for dated units, the vertically rearranged layout enables the archaeologist to visually derive temporal relations between units that are not stratigraphically related. This can also enhance the comparison and analysis of sequences from different sites.

The colour of a unit dated by a custom interval is mapped to the interval of the time model that completely contains the custom interval. In terms of Allen's Algebra:

Let X be the interval of the time model, and Y be the interval assigned to the unit, one of the following relations must be met so that both match: "Y during X", "Y starts X", "Y finishes X" or "X equals Y".

First, all intervals in the very right time frame are checked. If no match is found, elements in the middle time frame are checked next. Finally, the left time frame is checked. The colour is mapped to the matching interval of the time model. If no match is found, the unit is coloured white. To distinguish units assigned to an interval of the time model from ones with custom intervals, the latter are decorated with a thick border. Furthermore, undated units are coloured grey and have a solid thin border.

### 3.4. Clustering and grouping

Harris matrices can usually become very large and contain hundreds to thousands of units. To reduce visual complexity and provide an aggregated visualisation, we implemented tools so that the stratigraphic units can be hierarchically clustered and organised in parent groups (Figure 3, left). Additionally, the defined parent group can also be closed (Figure 3, right), further reducing complexity. Depending on the time intervals assigned to the children in the parent group, all children are combined into a single unit, which can extend over several time intervals.

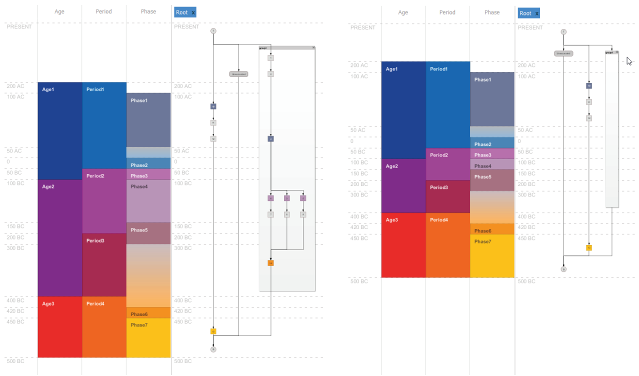

**Figure 3:** *Master graph with open (left) and closed (right) group*

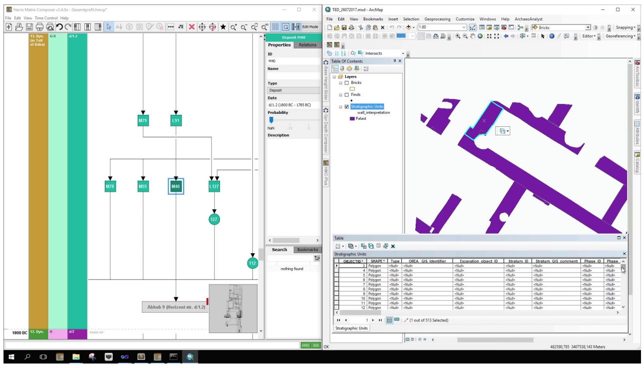

**Figure 4:** *Linked view between HMC+ and ArcMap 10.2*

To visualise dating information contained in closed groups, the height of the node representing a closed parent group is calculated. This is achieved by dividing the group node into two dummy nodes, one representing the upper time interval limit and one representing the lower time interval limit of all children. Now the time intervals for each node can be calculated individually. Finally, the height can then be derived from the difference between the lower and upper limit.

## 4. Interfacing the stratigraphic sequence to GIS

An interface to GIS makes the HMC+ to the ideal application for the management of digital archaeological data and thereby closes a gap in the software toolbox for archaeologist. In our case we linked the HMC+ with ESRI ArcMap 10.2. The HM is directly used to access data in the GIS for visualisation and analysis by selecting stratigraphic units in the HMC+. Each stratigraphic unit is associated with a shape in the GIS representing its boundary polygon. Linked views are established between both tools (Figure 4). This means that interactions in one tool immediately affects the other. For example, when selecting a unit in the HMC+ all the associated geographical and georeferenced objects are selected and highlighted or made visible in the GIS. The view in the GIS changes smoothly to display the selected objects optimally. On the other hand, a selected object in the GIS highlights the corresponding deposit or feature surface in the stratigraphic sequence displayed in HMC+.

Such linked views close the circle and push digital archaeology to the next level by tightly integrating all data channels (geographic, geometric, stratigraphic and chronologic), and make them available for scientific analysis in a consistent way.

## 5. Conclusions and outlook

The HM is a valuable concept for the representation of the unique stratification of an archaeological site investigated by a stratigraphic approach and in particular for the ordering of all individual units of stratification through time. The stratigraphic sequence or HM is the backbone for the further analysis of the excavation and the compilation of the description of the excavated and thus destroyed stratification, i.e. the stratigraphy of the excavated stratified volume.

The integrated approach for the digital documentation and visual analysis of a combination of topological (or stratigraphic) and temporal relations originating from an archaeological site presented here is an essential step for the virtual representations of archaeological stratifications to reliably support the most important task of an archaeologist which is the compilation of the archaeological stratigraphy of the excavated site.

The implicit chronologic sequence given by the HM becomes explicit as scientists are enabled to define a hierarchical time model and assign units of the HM to temporal intervals or provide exact dating. The system maintains a consistent visual representation, which means that a correct stratigraphic layout is preserved while units are aligned to intervals of the time model. Evaluation of several use cases such as the project "Tell el Daba" [KNMD16] showed that this combined visualisation makes the scientific analysis and interpretation more efficient and reliable. The new integrated framework will soon be available for the archaeological community, which will provide us with further valuable feedback and suggestions for improvements.

Our solution has two major limitations:

- Computational Time: Due to the consideration of temporal information, a local layout computation is not trivial. Therefore, the complete graph is recalculated for each new layout, which can lead to long computations for large HMs. An incremental approach has to be investigated to improve the overall computational time.
- Mental Map: Due to various optimisation steps inherent to the layout algorithm, such as the minimisation of crossing edges, it is difficult to maintain a consistent layout over time. Minor changes to the graph may lead to the horizontal exchange of vertical sequences. These location changes makes it difficult for the user to preserve a mental map of the graph.

Besides addressing these limitations in future work, we intend to advance the archaeological analysis by considering 3D reconstructions of stratigraphic layers, which today is the common practice. This allows to link HM units to 3D data sets and to derive the volume of deposits.

## References

[All83] ALLEN J. F.: Maintaining knowledge about temporal intervals. *Commun. ACM 26*, 11 (Nov. 1983), 832–843. URL: http://doi.acm.org/10.1145/182.358434, doi:10.1145/182.358434. 2, 3

[FH01] FLEISCHER R., HIRSCH C.: Graph drawing and its applications. In *Drawing Graphs*. Springer Berlin Heidelberg, 2001, pp. 1–22. doi:10.1007/3-540-44969-8_1. 3

[Han04] HANSON I. D.: The importance of stratigraphy in forensic investigation. *Geological Society, London, Special Publications 232*, 1 (2004), 39–47. doi:10.1144/gsl.sp.2004.232.01.06. 2

[Har89] HARRIS E. C.: *Principles of Archaeological Stratigraphy, Second Edition*. Academic Press, 1989. 1

[HMC18] Harris matrix composer, 2018. URL: http://www.harrismatrixcomposer.com/. 2

[IJ17] ICOVE J. H. G. A.: *Kirk's Fire Investigation (8th Edition)*. Pearson, 2017. 2

[JM04] JÜNGER M., MUTZEL P.: Technical foundations. In *Graph Drawing Software*. Springer Berlin Heidelberg, 2004, pp. 9–53. doi:10.1007/978-3-642-18638-7_2. 3

[KNMD16] KUCERA M., NEUBAUER W., MÜLLER S., DONEUS N.: The tell el daba archaeological information system: adding the fourth dimension to legacy datasets of long-term excavations (a puzzle in 4d), 04 2016. 4

[Neu07] NEUBAUER W.: Laser scanning and archaeology standard tool for 3d documentation of excavations. *GIM INTERNATIONAL 21*, 10 (2007), 14. 1, 2

[SM91] SUGIYAMA K., MISUE K.: Visualization of structural information: automatic drawing of compound digraphs. *IEEE Transactions on Systems, Man, and Cybernetics 21*, 4 (1991), 876–892. doi:10.1109/21.108304. 3

[STT81] SUGIYAMA K., TAGAWA S., TODA M.: Methods for visual understanding of hierarchical system structures. *IEEE Transactions on Systems, Man, and Cybernetics 11*, 2 (1981), 109–125. doi:10.1109/tsmc.1981.4308636. 3

[TN08] TRAXLER C., NEUBAUER W.: The harris matrix composer - a new tool to manage archaeological stratigraphy. 14th International Conference on Virtual Systems and. 1, 2

[yG18] YWORKS GMBH: yworks the diagramming company, 2004-2018. URL: https://www.yworks.com/. 2