Article:

Dye, T.S. and Buck, C.E. (2015) Archaeological sequence diagrams and Bayesian chronological models. Journal of Archaeological Science, 63. 84 - 93. ISSN 0305-4403

https://doi.org/10.1016/j.jas.2015.08.008

Title Archaeological Sequence Diagrams and Bayesian Chronological Models

Authors Thomas S. Dye[1] and Caitlin E. Buck

Affiliations University of Hawaiʻi at Mānoa and University of Sheffield

Address for correspondence 735 Bishop St., Suite 315, Honolulu, HI 96813

Email for correspondence tsd@tsdye.com

Approximate number of words 5071

---

[1] Corresponding author.

Abstract

This paper develops directed graph representations for a class of archaeological sequence diagrams, such as the Harris Matrix, that do not include information on duration. These "stratigraphic directed graphs" differ from previous software implementations of the Harris Matrix, which employ a mix of directed graph and other data structures and algorithms. A "chronological directed graph" to represent the relationships in a Bayesian chronological model that correspond to the possibilities inherent in a sequence diagram, and an algorithm to map a stratigraphic directed graph to a chronological directed graph are proposed and illustrated with an example. These results are intended to be a proof of concept for the design of a front-end for Bayesian calibration software that is based directly on the archaeological stratigrapher's identification of contexts, observations of stratigraphic relationships, inferences concerning parts of once-whole contexts, and selection of materials for radiocarbon dating.

Keywords Sequence diagram, Chronology, Directed graph, Bayesian radiocarbon calibration

Introduction

Advances in the methods and practice of radiocarbon dating in archaeology, sometimes characterized as revolutionary (Bayliss, 2009; Taylor, 1995; Linick et al., 1989), have worked generally to increase the precision of age estimates for archaeological events. A recent phase of this radiocarbon revolution has as its focus Bayesian calibration (Buck et al., 1996), which highlights the role of stratigraphic interpretation in the development of radiocarbon-based site chronologies. A key innovation of Bayesian calibration is its ability to integrate ancillary sources of chronological information with the information returned by the radiocarbon dating laboratory. In a typical archaeological application having to do with site chronology, records of the stratigraphic relationships of deposits and interfaces are a primary source of this ancillary information. Common sense indicates that a site chronology based on "the dates" *and* "the archaeology" is bound to be more reliable than one that relies only on one or the other (Bayliss, 2009, 127). The improvement yielded by Bayesian calibration has been demonstrated, perhaps most convincingly for the early Neolithic period of Southern Britain and Ireland where time-scales with resolutions that approach a human generation have been achieved (Bayliss et al., 2011). At Çatalhöyük, a Neolithic village in Anatolia, a basic goal of the Bayesian calibration is to provide "calendar date estimates for the construction, use, and disuse of the excavated buildings, in order to infer a structural narrative between buildings that are not stratigraphically related" (Bayliss et al., 2014, 69). Given that a typical house at Çatalhöyük was constructed, used, and disused over a period on the order of 60–145 years (Bayliss et al., 2014, 89), the ambitious goal of identifying contemporary houses from spatially separate parts of the site without the aid of dendrochronology (Towner, 2002) would have been wildly unrealistic prior to the development of AMS dating and Bayesian calibration.

The data requirements to achieve high precision estimates are sufficiently stringent that often specialists are sought to select samples for radiocarbon dating. The specialist works with a list of potential dating samples and a model of relative chronological relations yielded by stratigraphy, sometimes in the form of a sequence diagram such as the Harris Matrix (Harris, 1989) but more often in the form of profile drawings and excavation notes, to develop a chronological model that maximizes the value of the calibration results for interpretation. In effect, the specialist transforms one relative chronological model into another, moving from the stratigrapher's model expressed in terms of units of stratification, or contexts (Carver, 2005, 107), into the statistician's model expressed in terms of formal algebraic relationships between chronological phases.

This paper describes a transformation algorithm based on the theory of directed graphs that takes as its input a suitably structured sequence diagram and information on potential dating samples to produce a chronological model for use in Bayesian calibration. To demonstrate its utility in automating the creation of Bayesian chronological models, we apply the algorithm to Buildings 1 and 5 in the North Area at Çatalhöyük (Cessford, 2007d,c,b,a). This example represents a relatively rare situation where a detailed sequence diagram is published (Bayliss et al., 2014, Fig. 3.17) and dating specialists have carried out several Bayesian calibrations (Cessford et al., 2005; Bayliss et al., 2014).

Computing the Sequence Diagram

In archaeology, the term *sequence diagram* refers to a family of graphic displays designed to represent stratigraphic relationships (Carver, 2009, 276). Perhaps the most widely used sequence diagram is produced by the Harris Matrix, which is described by its creator as a method by which the order of the deposition of the layers and the creation of feature interfaces through the course of time on an archaeological site can be diagrammatically expressed in very simple terms (Harris, 1989, 34). This focus on the order of deposition to the exclusion of other attributes distinguishes the Harris Matrix from sequence diagrams which augment the order of deposition with information about duration (Dalland, 1984; Carver, 1979), and it is this sense in which sequence diagram is used here.

Since the transformation algorithm we propose is based on the theory of directed graphs, the sequence diagram used as input must be capable of representation as a directed acyclic graph, or DAG, which can be manipulated programatically. A DAG conceptualizes the stratigraphic structure of an archaeological sequence as chronological relationships on a set of depositional and interfacial contexts. A directed graph consists of one or more of a finite set of nodes and zero or more connections between ordered pairs of distinct nodes, each of which defines an arc (Harary et al., 1965). In the case of archaeological stratigraphy, an archaeological context is represented as a node and a stratigraphic relationship between two contexts is represented by an arc.

Available Harris Matrix software packages are closed-source and do not permit programmatic access to the DAG representation, so it proved necessary to develop the open-source software package, `hm`, to achieve this goal (provided as supplementary material). Although computer programmers quickly recognized that the sequence of observed stratigraphic relationships at the heart of the sequence diagram can be represented as a DAG (Ryan, 1988; Herzog, 1993; Herzog and Scollar, 1991), the display conventions of the Harris Matrix are tied to the layout of paper forms developed in the 1970s (Harris, 1989, 34) and these conventions introduce complexities that can not be represented by a DAG. Thus, the `hm` software abandons certain display conventions of the Harris Matrix in order to preserve a pure DAG representation of the sequence diagram.

The following sections compare and contrast DAG and Harris Matrix representations of the sequence diagram and present the data inputs to the `hm` software as tables that define entities in a relational database (fig. 1). The first three sections consider the relationships between contexts recognized by the Harris Matrix— i) no direct stratigraphic relationship, or context identity, ii) an observed relationship of superposition, and iii) parts of a once-whole context —in turn, as steps in the construction of a sequence diagram. This is followed by a consideration of periods and phases, which are conceptually similar interpretive constructs.

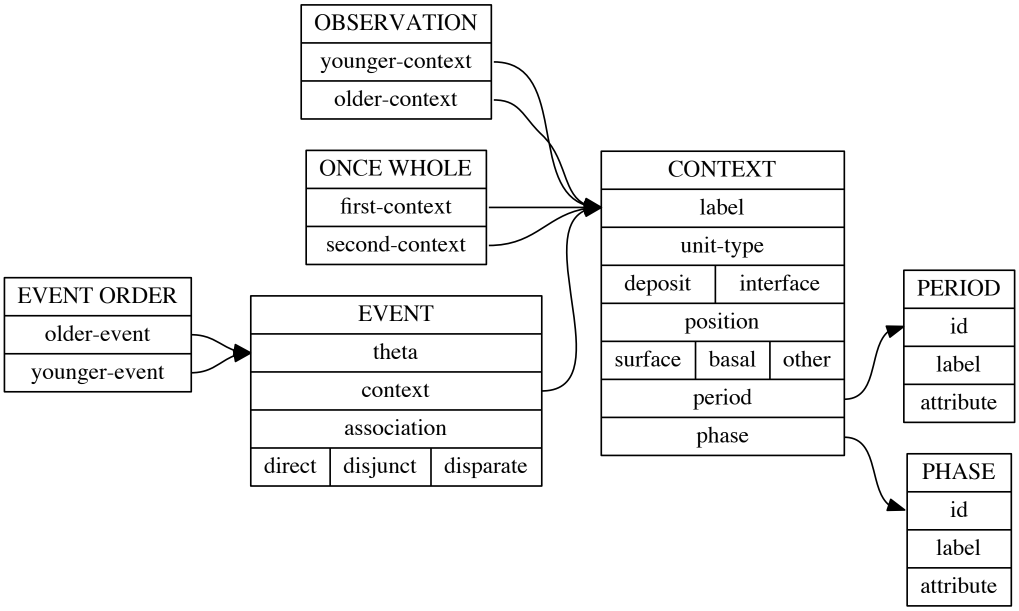
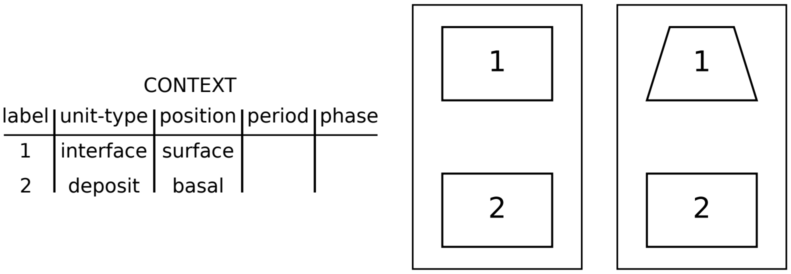


Figure 1: Relational database design for the seven tables of information used to construct stratigraphic and chronological directed graphs. Note that table names are uppercase, column names are lowercase, and divided entries define the domain of the column whose name is directly above, e.g., the `unit-type` column in the context table contains one of the two values `deposit` and `interface`.

Identification of Contexts

Archaeologists commonly identify five types of context: deposits, horizontal feature interfaces, vertical feature interfaces, upstanding layer interfaces, and horizontal layer interfaces. The Harris Matrix was designed, in part, to ensure that all of the contexts identified at a site are included in the sequence (Roskams, 2001, 157) and to replace the previous archaeological practice of recording contexts and their relationships with section drawings, which typically take in only some small fraction of the contexts identified at a site (Bibby, 1993, 108).

In practice, the archaeologist working with a printed Harris Matrix sheet draws up a list of identified depositional and feature interface contexts, then writes each context identifier in a rectangular box on the grid. Contexts close to one another in space are placed in rectangular boxes close to one another on the grid and the vertical position is chosen to reflect the context's position in the stratigraphic sequence, with surficial contexts placed near the top of the diagram and basal contexts placed near the bottom. At this stage the Harris Matrix consists of rectangular boxes with context identifiers within them, and the rectangular boxes are not yet connected to one another (fig. 2, *center*).



CONTEXT

label	unit-type	position	period	phase
1	interface	surface		
2	deposit	basal		

Figure 2: Initial stage in construction of a sequence diagram consisting of an interface, context 1, and a deposit, context 2: *left*, a five-column context table that records information about contexts (see fig. 1); *center*, Harris Matrix; *right*, directed graph.

By convention, horizontal layer interfaces are not represented in the Harris Matrix because they are considered to have "the same stratigraphic relationships as the deposits and are recorded as an integral part of the layers" (Harris, 1989, 54). This practice appears to be deeply ingrained in the archaeological community, but it is problematic from the point of view of relative chronology (Clark, 2000, 103). Treating the layer interface as an integral part of the depositional context beneath it ignores the possibility that it represents a unit of time, either because the surface it represents was deflated by erosion, exposing old deposits, or because the surface itself was open for some time. The failure to record layer interfaces potentially introduces hiatuses into the chronological model. A hiatus-free sequence diagram (and thus the associated directed graph) exhibits a particular structure with alternating interfacial and depositional contexts. In contrast, conventional stratigraphic practice places deposits in a relationship of direct superposition across unrecorded layer interfaces. Of course, archaeologists who use the Harris Matrix recognize the unrecorded layer interfaces and these are brought back into the analysis at a later stage, when periods are identified (Harris, 1989, Fig. 25). It is at this late analytic stage that the definition of a period boundary as an interface and its specification in the Harris Matrix as a mix of interfaces and deposits is reconciled (Harris, 1989, 67–68).

Because the representation of a directed graph is not constrained by the conventions of the Harris Matrix, the shapes of nodes can express the fundamental distinction between depositional and interfacial contexts. The convention adopted here uses a rectangular box, similar to the symbol used in a Harris Matrix, when `unit-type` is set to `deposit` and a trapezium when `unit-type` is set to `interface` (fig. 2, *right*).

Observed Stratigraphic Relationships

The next step in construction of the sequence diagram is to indicate observed stratigraphic relationships. In practice, the stratigrapher records observed relationships in a two-column table, where one column contains the identifiers of the younger contexts that assume a superior position in the observed stratigraphic relationship and the other column contains the identifiers of the older contexts that assume an inferior position in the observed stratigraphic relationship (fig. 1). For each row of the table, the stratigrapher identifies on the sequence diagram the rectangular box that represents the younger context and searches below it for the rectangular box that represents the older context. An orthogonal line is then drawn from the bottom of the rectangular box representing the younger context to the top of the rectangular box representing the older context (fig. 3, *center*).

OBSERVATION

younger-context	older-context
1	2

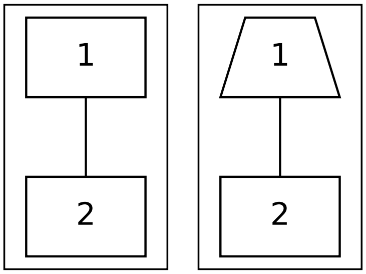


Figure 3: The sequence diagram after stratigraphic relationships are indicated with vertical lines: *left*, a two-column observation table that records the stratigraphic relationship between contexts 1 and 2 (see fig. 1); *center*, a Harris Matrix showing a younger interface, context 1, overlying an older deposit, context 2; *right*, a directed graph showing a younger interface, context 1, overlying an older deposit, context 2.

The directed graph uses the same table of observed stratigraphic relationships that the stratigrapher uses to draw the Harris Matrix. It is easy to see that each row of the table (fig. 3, *left*) represents an ordered pair of nodes, which in the theory of directed graphs defines an arc. The ordering is given by the stratigraphic relationship of the nodes; the younger context is by convention designated the start node of the arc and the older context the end node. It is customary to represent the arcs in a directed graph as arrows, with an arrowhead at the end of each arc to indicate direction. However, the Harris Matrix convention that uses a plain line and indicates direction by vertical position, such that a younger context appears above an older context with which it shares a stratigraphic relationship, is appreciated by archaeologists who see in it the physical relationship of the contexts when viewed in section. Thus, the directed graphs presented here adopt this convention and draw arcs as lines rather than arrows (fig. 3, *right*).

At this stage in its construction, the Harris Matrix is a partial order, or poset (Orton, 1980, 67). The stratigraphic relationships that it records are irreflexive, because an archaeological context cannot be stratigraphically superior or inferior to itself, asymmetrical because a context that is stratigraphically superior to another context cannot be stratigraphically inferior to it, and transitive because, given three contexts, *1*, *2*, and *3*, if *1* is stratigraphically superior to *2*, and *2* is stratigraphically superior to *3*, then *1* is stratigraphically superior to *3*.

Parts of Once-Whole Contexts

In the Harris Matrix, pairs of contexts inferred to have been part of a once-whole context are connected with two horizontal lines to indicate this relationship (fig. 4, *bottom left*). The information needed for this step is a table with two columns, where each row represents an inference that the two contexts in it are parts of a once-whole context (fig. 1). Parts of a once-whole context describe a symmetrical relation that is transitive; this type of relation is outside the theory of directed graphs. Parts of a once-whole context can be treated in two ways by a directed graph. In the first, the directed graph is used to model only observations of stratigraphic relationships; inferred parts of a once-whole context can be plotted at the same vertical level of the sequence diagram, but stratigraphic relationships implied by the inference of once-wholeness are not taken into account (fig. 4, *top right*). In the second, the inference of once-wholeness is assumed to be true and parts of a once-whole context are treated as a single context (fig. 4, *bottom right*). Thus, the Harris Matrix displays in a single sequence diagram observations of stratigraphic relationships and inferences about parts of once-whole contexts; two directed graphs are required to show the same information.

Stratigraphic Periods and Phases

The terms “period” and “phase” are defined variously and sometimes interchangeably by archaeologists. For the Harris Matrix, a “phase” groups contexts of similar age, and a “period” groups phases of similar age, yielding a nested series of time intervals (Harris, 1989, 158). Defined in this way, both phases and periods are interpretive constructs that are typically formulated with both stratigraphic and non-stratigraphic information. Because “phase” is also used to describe Bayesian chronological models, here we use the term “stratigraphic phase” to refer to a group of contexts, and the term “chronological phase” to refer to a time period in a chronological model.

Alternative ways to represent periods and stratigraphic phases can be illustrated using a stratigraphic profile drawing developed by Harris (1989, Fig. 12a) and adapted here (fig. 5).

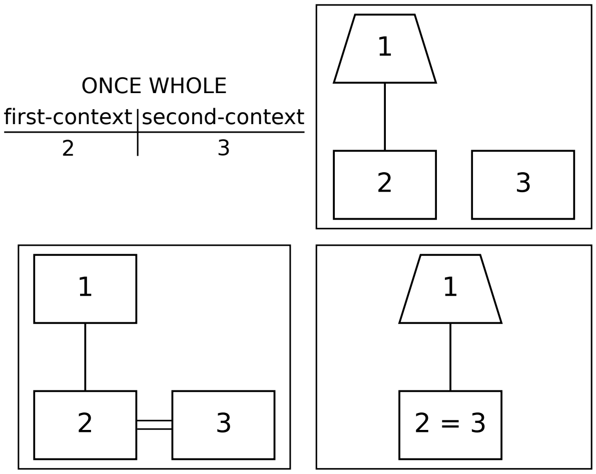


Figure 4: Three graphical representations of parts of a once-whole context: *top left*, two-column once-whole table recording the inference that contexts 2 and 3 are parts of a once-whole context (see fig. 1); *bottom left*, the Harris Matrix connects contexts 2 and 3 with two horizontal lines; *top right*, a directed graph representation of the observed relationships of superposition places contexts 2 and 3 at the same level, but does not make explicit the inferred stratigraphic relationship between contexts 1 and 3; *bottom right*, a directed graph representation of the sequence diagram where the inferred relationship between contexts 2 and 3 as parts of a once-whole context is assumed to be true and the contexts have been merged and labeled "2 = 3".

The Harris Matrix displays periods and stratigraphic phases in the same way, by horizontal lines drawn across the diagram (fig. 6, *left*). In contrast, the directed graph convention displays periods and stratigraphic phases by altering the graphic attributes of nodes (fig. 6, *right*).

Structure of a Bayesian Chronological Model

The chronological model now widely used in Bayesian chronology construction comprises entities different than those of an archaeological sequence diagram. The basic entity of a sequence diagram is a stratigraphic context; a Bayesian chronological model comprises directly-dated events and the start and end dates of one or more chronological phases. The start and end dates of a chronological phase typically map directly to an archaeological context, and so in this paper we will assume that no additional information is needed to represent them beyond that which is available from the stratigraphic directed graph.

Within software such as `hm`, it is convenient to capture the information about dated events

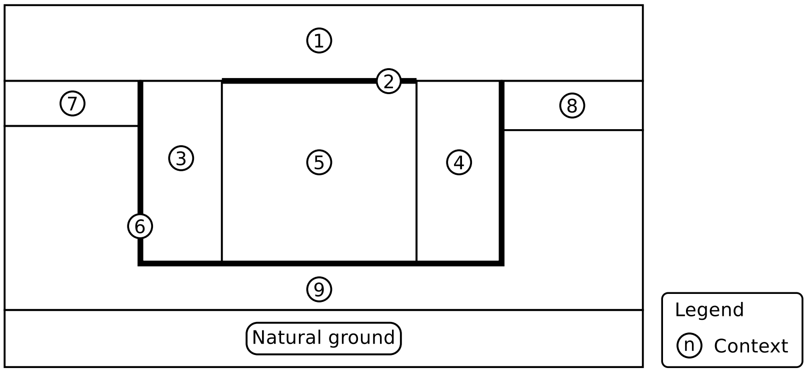


Figure 5: Illustrative stratigraphic profile drawing. Adapted from Harris (1989, Fig. 12a).

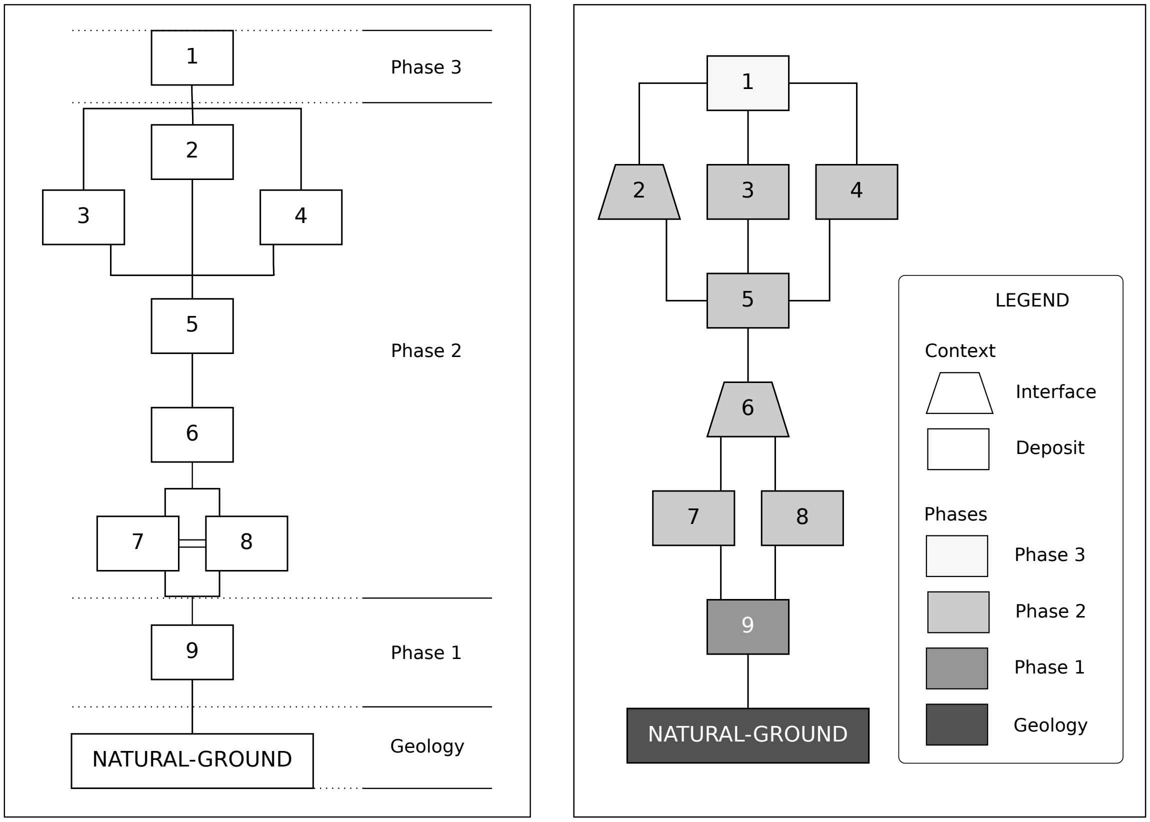


Figure 6: Hypothetical phasing of an example sequence developed by Harris (see fig. 5): *left*, the Harris Matrix representation, after Harris (1989, Fig. 12c); *right*, directed graph representation with nodes shaded to indicate phases.

in two tables. An "event table" associates a directly-dated archaeological event with its archaeological context (fig. 1) and indicates whether the event is directly associated with the context, is older than the context and thus disjunct, or is younger than the context and thus disparate (Dean, 1978). An "event order table" records information on the relative ages of archaeological events associated with the same context (fig. 1).

One difference between a Bayesian chronological model and an archaeological sequence diagram is that the Bayesian chronological model may include relationships that cannot be expressed by stratigraphy. An illustration recognizes three possible relationships between two chronological phases where one is older than the other (fig. 7). Only two of these relationships can be represented stratigraphically.

- One chronological phase can be older than the other such that the end date for the older chronological phase is older than the start date for the younger chronological phase (fig. 7, *left*). This relationship, where a time interval separates two chronological phases, arises in archaeological stratigraphy when two contexts are found on the same line of a (possibly multi-linear) sequence but are separated by one or more contexts. This relationship is relatively common in practical Bayesian chronological models. Contexts that lack dating material are typically ignored in a Bayesian chronological model.

- One chronological phase can be older than the other such that the end date for the older chronological phase is the same age as the start date for the younger chronological phase (fig. 7, *middle*). This abutting relationship describes the relationship of superposition that archaeologists typically observe in the field.

- One chronological phase can be older than the other such that the end date for the older chronological phase is younger than the start date for the younger chronological phase (fig. 7, *right*). This overlapping relationship cannot be determined solely on stratigraphic grounds because the two contexts must be from different lines of a multi-linear stratigraphic sequence. Other information, perhaps having to do with the content of the contexts, is required to posit this kind of relationship (Triggs, 1993).

Another difference between a Bayesian chronological model and an archaeological sequence diagram is that the archaeological sequence diagram is concerned only with relationships between archaeological contexts, but the chronological model includes relationships among a variety of different entities, including early phase boundaries, late phase boundaries, and dated events. In addition, the notation for recording relationships between phase boundaries must distinguish between phase boundaries that share the same calendar age and phase boundaries that are separated in time. For example, depositional context i, within which a single event, e, was identified and dated might be represented by the chronological model as $\alpha_i > \theta_e > \beta_i$, where α_i and β_i are the start and end dates, respectively, of chronological phase i, θ_e represents the calendar age of event e, and $>$ means "is

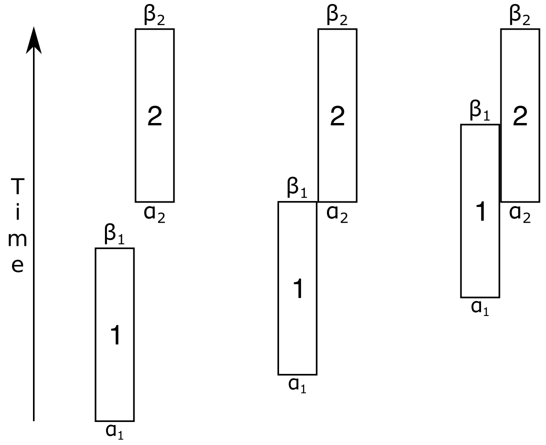


Figure 7: Schematic representation of three possible relationships between older and younger chronological phases: *left*, chronological phases 1 and 2 separated, $\beta_1 > \alpha_2$; *middle*, chronological phases 1 and 2 abutting, $\beta_1 = \alpha_2$; *right*, chronological phases 1 and 2 overlapping, $\beta_1 < \alpha_2$. Adapted from Buck et al. (1996, Fig. 9.8).

older than". Alternatively, this simple chronological model can be represented as a directed graph (fig. 8, *left*), where vertical position represents relative age, similar to the convention used in directed graphs of archaeological sequences.

Mapping a Sequence Diagram to a Chronological Model

Given a directed graph of a hiatus-free archaeological sequence from which transitive relationships have been removed, it is possible to construct a Bayesian chronological model by combining the relative chronological information in the directed graph of the archaeological sequence diagram with the potentially dated events. Recall that a directed graph consists of a finite set of nodes and a collection of ordered pairs of distinct nodes, the connection between any pair of which is called an arc. Two nodes connected by an arc are said to be *adjacent*; the start node of the arc is *adjacent to* the end node, and the end node is *adjacent from* the start node. The *outdegree* of a node is the number of nodes adjacent from it, and the *indegree* of the node is the number of nodes adjacent to it. A *walk* in a directed graph is an alternating sequence of nodes and arcs, and a *path* is a walk in which all nodes are distinct. If there is a path from node u to node v, then node v is *reachable* from node u. The directed graph concept of reachability can be used to determine whether two contexts are on the same line of a possibly multilinear sequence diagram. If, for two archaeological contexts, x and y, x is reachable from y or y is reachable from x, then x and y are on the same line of

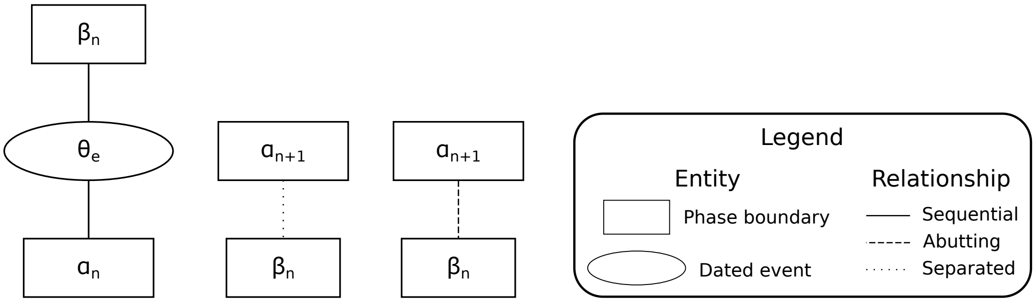


Figure 8: Entities and relationships of Bayesian chronological models represented as directed graphs: *left*, a chronological phase with a single dated event; *middle*, relationship between boundary parameters of separated chronological phases; *right*, relationship between boundary parameters of abutting chronological phases.

the sequence diagram. Conversely, if x is not reachable from y and y is not reachable from x, then x and y are on different lines of a multi-linear sequence diagram.

For two archaeological contexts, x and y, on the same line of an hiatus-free sequence diagram such that y is reachable from x, the directed graph concept of adjacency can be used to distinguish an abutting chronological relationship, where x is adjacent to y, from a separated relationship, where x is not adjacent to y. These relationships are illustrated in Figure 9, which categorizes contexts according to their chronological relationship to Context 4 using a directed graph that includes contexts and their observed stratigraphic relationships (fig. 9, *center*) and one that augments this information with inferences about once-whole contexts (fig. 9, *right*). These graphs indicate that directed graph representations of an archaeological sequence contain the information needed to construct a Bayesian chronological model.

The maximal chronological directed graph is obtained by adding to the stratigraphic directed graph extra nodes and arcs to represent the information in the event table and the event order table. Since the number of contexts with potentially dated events is typically much smaller than the number of undated ones, however, an algorithmic version of this approach would not closely mirror what those constructing Bayesian models do at present. A six step algorithm can, however, be used to construct the minimal chronological directed graph (and hence chronological model) from the directed graph of the archaeological sequence and the two tables of potentially dated event information, as follows.

Suppose the set of all contexts in our stratigraphic directed graph is $\mathcal{C}$ and that the subset of those with potentially dated events is $\mathcal{D}$. The number of elements in $\mathcal{D}$, $\#\mathcal{D}$, is typically much smaller than the number in $\mathcal{C}$ since relatively few contexts from the excavation contain potentially dated finds. The set of potentially dated events (i.e. events in the event table) is then $\mathcal{E}$ with individual elements $\{e_1, e_2, \cdots, e_E\}$, where $E = \#\mathcal{E}$. Each member of

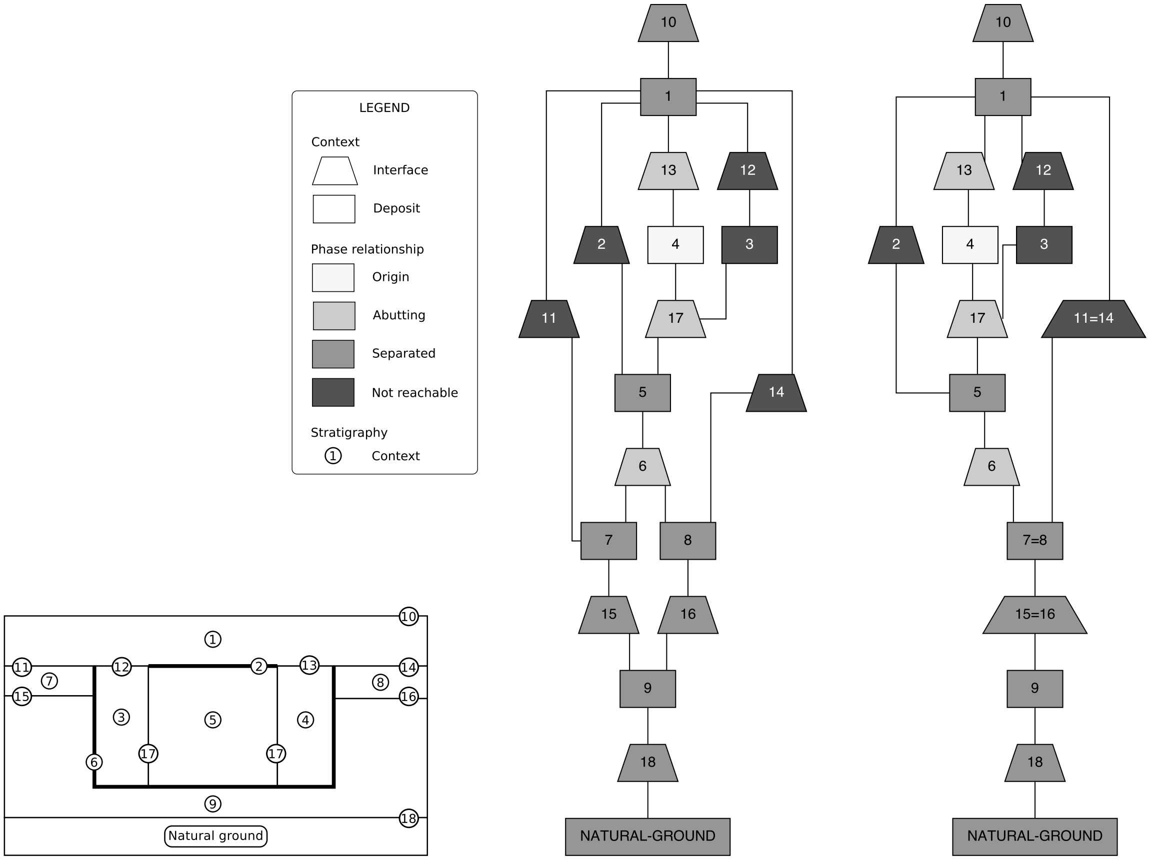


Figure 9: A hiatus-free sequence diagram with contexts shaded according to their chronological relationship to Context 4: *left*, the stratigraphic profile after Harris (1989, Fig. 12), with layer interfaces numbered 10–18 (cf. Fig. 5); *center*, a directed graph representation of the sequence diagram depicting observed relationships of superposition; *right*, a directed graph representation of the sequence diagram in which inferences of once-whole contexts are assumed to be true.

$\mathcal{E}$ was excavated from a context and so is associated with one and only one member of $\mathcal{D} = \{d_1, d_2, \cdots, d_D\}$, where $D = \#\mathcal{D}$.

1. For each member of $\mathcal{D}$, d_i, add two nodes to the chronological directed graph, one for the early boundary date, α_{d_i}, and the other for the late boundary date, β_{d_i}.
2. For each member, e_j, of $\mathcal{E}$ add one node, θ_{e_j}, to the chronological directed graph to represent its absolute date.
3. For each row in the event order table add an arc from the younger node to the older node.
4. For each row, $j = 1, 2, \cdots, E$, of the event table (associated with archaeological context d_i and event with absolute date θ_{e_j}):
    a) if the indegree of θ_{e_j} is 0 (and association is not equal to "disjunct") add an arc from α_{d_i} to θ_{e_j} and assign it a value of 0;
    b) if the outdegree of θ_{e_j} is 0 (and association is not equal to "disparate") add an arc from θ_{e_j} to β_{d_i} and assign it a value of 0.
5. For each pair (d_l, d_m) of archaeological contexts in the event table:
    a) if d_l is reachable from d_m in the directed graph of the archaeological sequence, add an arc from β_{d_l} to α_{d_m} in the chronological directed graph;
    b) if context d_m is adjacent to context d_l in the directed graph of the archaeological sequence, assign the arc from β_{d_l} to α_{d_m} in the chronological directed graph a value of 1, else assign it a value of 2;
    c) If context d_l is reachable from context d_m in the directed graph of the archaeological sequence, add an arc from β_{d_m} to α_{d_l} in the chronological directed graph;
    d) if context d_l is adjacent to context d_m in the directed graph of the archaeological sequence, assign the arc from β_{d_m} to α_{d_l} in the chronological directed graph a value of 1, else assign it a value of 2.
6. Perform transitive reduction.

Discussion

At present, it appears to be the case that no archaeologists build their chronological models using formal algorithms. Instead they apply their expert judgment, selecting features from the stratigraphic record to include in the model on whatever basis they choose and justify their decisions in prose in the resulting publication. Such an approach may well lead archaeologists to learn all they wish to from the chronological evidence available, but it would be hard to demonstrate that and few authors at present even discuss the impact of their choice of chronological model on the results obtained.

An example where the authors do discuss the impact of model choice is the work undertaken to establish the chronology of Buildings 1 and 5 in the North Area excavations at Çatalhöyük (Cessford et al., 2005; Bayliss et al., 2014). The initial work was exploratory in nature, with one goal "to determine which types of material and/or context provide good dating evidence" (Cessford et al., 2005, 84). The reliability of each dated sample was ranked as "low" where "there is a direct stratigraphic relationship between determinations that contradicts the relationship between the ages of the two determinations" (Cessford et al., 2005, 76), "high" where the sample comes from "a consistently dated stratigraphic sequence" (Cessford et al., 2005, 76) or where it is "short lived material from a context with a low probability of residuality" (Cessford et al., 2005, 76), or "medium" otherwise (Supplementary Material Table S1). Where possible, contradictions were resolved with reference to four of the five age determinations from Context 1332+[2] in Building 1, a "deliberately-placed deposit of lentils which represents a single year's harvest of a short-lived species that was purposefully burnt" (Cessford et al., 2005, 86). Context 1332+ has a direct stratigraphic relationship with all of the contexts excavated from Building 5, which underlies Building 1, but its age relative to most of the contexts from Building 1 cannot be determined (fig. 10).

Since the full sequence diagram for Buildings 1 and 5 is too large to reproduce here and given its pivotal role in the interpretation of the chronology of both buildings, we focus our illustration on Context 1332+ and those closest to it stratigraphically. However, the full sequence diagram and the chronological models derived via our algorithm are provided in the Supplementary Material.

A directed graph representation of the chronological model implied by the exploratory analysis accepts the assumption that each dated sample is associated with the context from which it was collected (fig. 11). The chronological model indicates that none of the related contexts superior to Context 1332+ in Building 1 were dated. Of the six dated contexts that are stratigraphically related to Context 1332+, five are from Building 1 and one, Context 3810+, is from Building 5. Thus, potential contradictions could be worked out with direct reference to the lentil deposit for a small subset of the dated contexts.

Carrying through the exploratory approach, Cessford et al. rejected the age determination for one of the lentils, θ_{31}, as inconsistent with the other four age determinations on lentils from Context 1332+, θ_{29}, θ_{30}, θ_{32}, and θ_{33}. Two dates on animal bone, θ_{42} from Context 1295a+ and θ_{24} from Context 1456, were assigned medium reliability because they were older than botanical material from the same deposits and the four lentils (Cessford et al., 2005, 88). As can be seen in Supplementary Material Figure S1, these comparisons with the lentils are not based on stratigraphic relationships; Contexts 1295a+ and 1456 are not reachable from Context 1332+ and their relative ages cannot be determined on stratigraphic

---

[2] It was frequently the case that a single context was assigned two or more field numbers. These field numbers were carried through the analysis and appear on the published Harris matrix for the excavation (Bayliss et al., 2014, Figure 3.17). The convention adopted here typically uses the first field number assigned to a context and indicates multiple field numbers for a single context by appending a "+" to the field number.

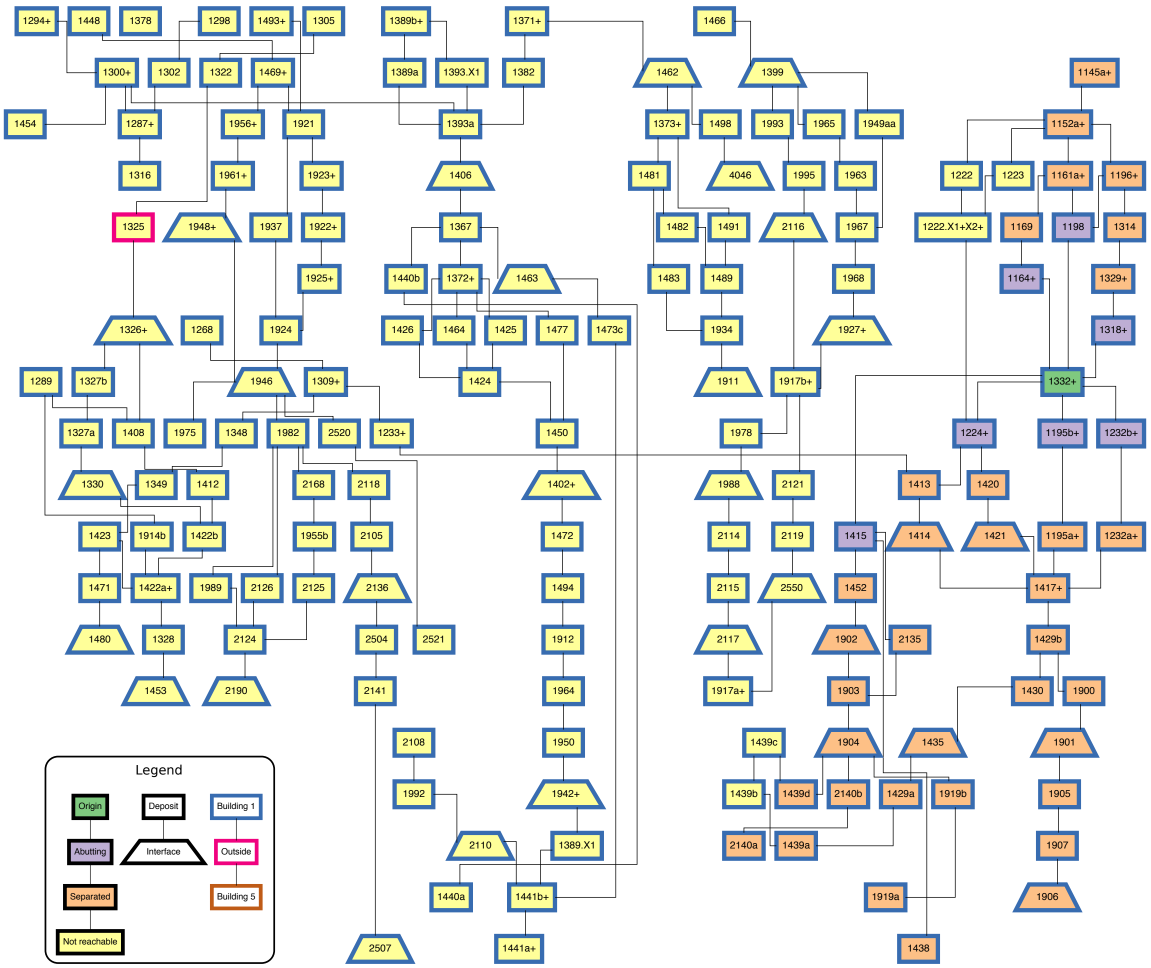


Figure 10: A portion of the sequence diagram for Buildings 1 and 5 of the North Area excavations at Çatalhöyük showing Context 1332+ in Building 1, adjacent and reachable contexts whose ages relative to Context 1332+ are known, and unreachable contexts whose ages relative to Context 1332+ can not be determined stratigraphically. Note that the majority of the contexts shown on the diagram are deposits and that interfacial contexts are comparatively rare. The full sequence diagram, of which this is a part, is available as Supplementary Material Figure S1.

grounds. Instead, the comparison appears to be made on the basis of "the division of the site into phases" (Cessford et al., 2005, 65), and thus on inferences rather than direct observations. Similarly, six dates on human bone were considered to be "in agreement with the stratigraphic sequence and the determinations from the lentils" (Cessford et al., 2005,

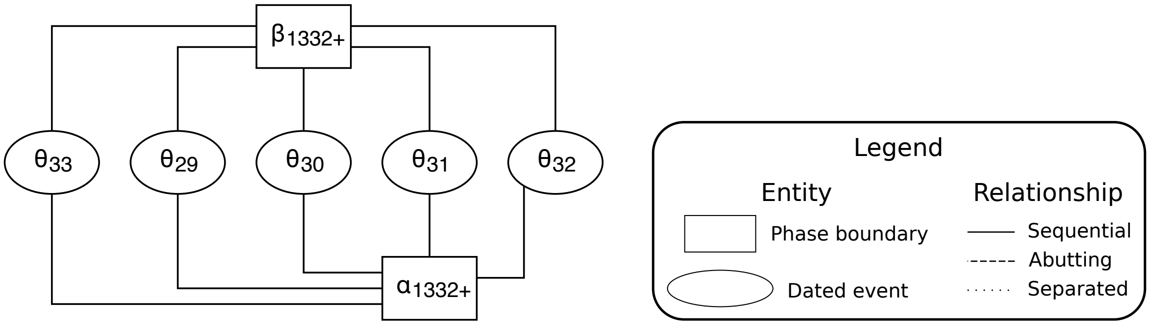


Figure 11: Representation of the dated lentils from Context 1332+ on a chronological model for determining which types of material and/or context provide good dating evidence using the dated samples reported by Cessford et al. (2005, Table 4.10) and the sequence diagram for the North Area excavations (fig. 10). The full chronological model is available as Supplementary Material Figure S2.

87), however five of these dates, θ_{49-53}, have no stratigraphic relationship to the lentils, and these comparisons also appear to be a result of phasing. One age determination, θ_{48} from Context 2519, is stratigraphically inferior to the lentils and so directly comparable.

Subsequently, dates on human bone and antler processed at the Oxford Radiocarbon Accelerator Unit between 2000 and 2002 were shown to be incorrect due to a technical problem. When re-dated, the bone and antler samples from Çatalhöyük, including the six dates on human bone, were determined to be 50–150 BP younger than the original measurements (Bayliss et al., 2014, 79). In particular, θ_7, which replaced θ_{48} from Context 2529, stratigraphically inferior to the lentils, returned a date younger than the four lentils, but older than the lentil that was previously rejected. Accordingly, the four lentils previously determined to represent the true age of the lentil deposits were interpreted as residual, and the lentil previously believed to be a statistical outlier was accepted as dating the true age of the deposit. This circumstance, and a comprehensive reevaluation of the suitability of the dated sample materials based largely on experience gained subsequent to the original exploratory dating project (Bayliss et al., 2014, 81–88), resulted in a different chronological model, one in which a large proportion of the dated samples are *termini post quem* for the end date of the context from which they were collected but have no relationship to the start date (fig. 12). These “dangling θ ’s” graphically illustrate the substantial challenges posed by residuality for the ambitious dating project at Çatalhöyük.

Conclusions

Directed acyclic graphs are already in widespread use in a number of disciplines in which, for reasons of practicality or logic, a collection of tasks or ideas must be ordered into

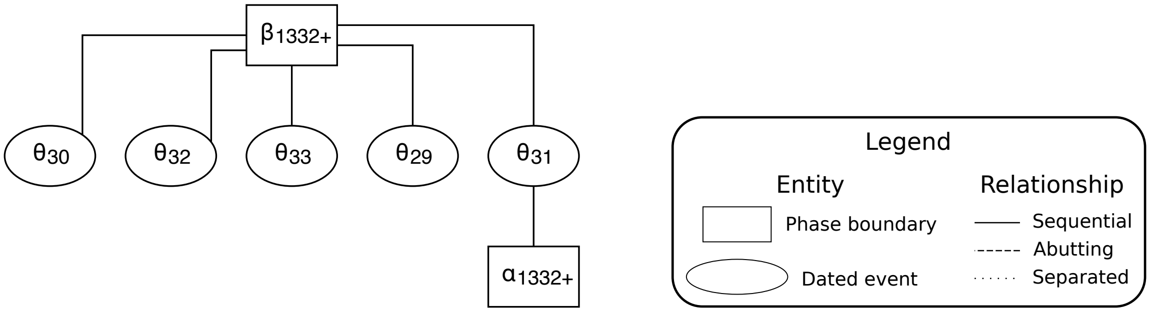


Figure 12: Representation of the dated lentils from Context 1332+ on a revised chronological model using the dated samples reported by Bayliss et al. (2014, Table 3.2) and the sequence diagram for the North Area excavations (fig. 10). The full revised chronological model is available as Supplementary Material Figure S3.

a sequence. Many well established algorithms now exist for performing inference on ideas that are represented as DAGs including, for example, the Markov chain Monte Carlo (MCMC) algorithms now so widely used in Bayesian inference in general and in Bayesian chronological modelling in particular.

Like many other statistical models, Bayesian chronological models are hierarchical in nature, with calendar ages of individual samples, linked sequentially to those for contexts, phases, structures, and so on. Such models have for many years been represented as DAGs both in publications (Parent and Rivot, 2013; King et al., 2010) and in software tools. Of the latter, the general purpose Bayesian inference environment known as WinBUGS (Lunn et al., 2000) – one of the first to become widely used – allows users the choice to define their model via a DAG from which the software generates the Bayesian model automatically.

One natural future use of the construction of chronological directed graphs from stratigraphic ones would thus be as a front-end to Bayesian chronological modelling software. Users could then develop a plethora of chronological directed graphs (based on automated algorithms, expert judgment, or both), estimate the parameters of the resulting models given real or simulated data, compare the resulting chronologies and even conduct formal model choice to establish which model best fits the currently available data.

Prototype software for creating and illustrating both stratigraphic and chronological directed graphs was developed to carry out the analyses in this paper.[3] The software establishes that the conversion from archaeological sequence diagram to a Bayesian chronological model can be made entirely rule-based and thus relatively straightforward. However, if others wish to benefit from these developments, and particularly if the automated generation of chronological directed graphs from stratigraphic ones is seen as beneficial, then

---

[3] The free and open-source Common Lisp software can be accessed at `http://tsdye.github.io/harris-matrix/`.

more work is needed. The next phase of this project will thus involve close collaboration with those who code Bayesian chronological modeling software with a view to providing a directed graph front-end that will offer a more intuitive way for archaeologists to build chronological models than such software offers at present and, ultimately, allow systematic exploration of the impact of different models on the chronological inferences made.

Acknowledgments

The authors thank Alex Bayliss for suggesting the example of Çatalhöyük Buildings 1 and 5 and for providing guidance during our analysis; Craig Cessford for clarifying conventions used in the representation of the Harris Matrix for Buildings 1 and 5 at Çatalhöyük; Julian Richards, Keith May, and Kieron Niven for advice on data standards and assistance searching the ADS archives; Eric Schulte for the `graph.lisp` library and for patient help during development of the `hm` software; and two anonymous reviewers for recommending that the paper include a real-world example. Any errors are the authors'.

References

Bayliss, A., 2009. Rolling out revolution: Using radiocarbon dating in archaeology. Radiocarbon 51, 123–147.

Bayliss, A., Farid, S., Higham, T., 2014. Time will tell: Practising Bayesian chronological modeling on the East Mound, in: Hodder, I. (Ed.), Çatalhöyük Excavations: The 2000–2008 Seasons. British Institute at Ankara and Cotsen Institute of Archaeology Press, London and Los Angeles. number 46 in British Institute at Ankara Monograph, pp. 53–90. Çatalhöyük Research Project Volume 7.

Bayliss, A., van der Plicht, J., Bronk Ramsey, C., McCormac, G., Healy, F., Whittle, A., 2011. Towards generational time-scales: the quantitative interpretation of archaeological chronologies, in: Gathering Time: Dating the Early Neolithic Enclosures of Southern Britain and Ireland. Oxbow Books, Oxford. chapter 2, pp. 17–59.

Bibby, D.I., 1993. Building stratigraphic sequences on excavations: an example from Konstanz, Germany, in: Harris et al. (1993). chapter 7. pp. 104–121.

Buck, C.E., Cavanagh, W.G., Litton, C.D., 1996. Bayesian Approach to Interpreting Archaeological Data. Statistics in Practice, John Wiley & Sons, Chichester, UK.

Carver, M., 1979. Three Saxo-Norman tenements in Durham City. Medieval Archaeology 23, 1–80.

Carver, M., 2005. Key ideas in excavation, in: Renfrew, C., Bahn, P. (Eds.), Archaeology: The Key Concepts. Routledge, London, pp. 106–110.

Carver, M., 2009. Archaeological Investigation. Routledge, New York.

Cessford, C., 2007a. Building 1, in: Hodder (2007). pp. 405–530. Çatalhöyük Research Project Volume 3.

Cessford, C., 2007b. Building 5, in: Hodder (2007). pp. 361–403. Çatalhöyük Research Project Volume 3.

Cessford, C., 2007c. History of excavation of Buildings 1 and 5 and summary of phases, in: Hodder (2007). pp. 345–360. Çatalhöyük Research Project Volume 3.

Cessford, C., 2007d. Overall discussion of Buildings 1 and 5, in: Hodder (2007). pp. 531–549. Çatalhöyük Research Project Volume 3.

Cessford, C., Blumbach, M., Akoğlu, H.G., Higham, T., Kuniholm, P.I., Manning, S.W., Newton, M.W., Ozbakan, M., Ozer, A.M., 2005. Absolute dating at Çatalhöyük, in: Hodder, I. (Ed.), Inhabiting Çatalhöyük: Reports from the 1995–1999 Seasons. McDonald

Institute for Archaeological Research and British Institute of Archaeology at Ankara, Cambridge and London. number 38 in British Institute at Ankara Monograph, pp. 65–99. Çatalhöyük Research Project Volume 4.

Clark, P., 2000. Negative features and interfaces, in: Roskams, S. (Ed.), Interpreting Stratigraphy: Site Evaluation, Recording Procedures and Stratigraphic Analysis. Archaeopress, Oxford. number 910 in BAR International Series. chapter 11, pp. 103–105.

Dalland, M., 1984. A procedure for use in stratigraphic analysis. Scottish Archaeological Review 3, 116–127.

Dean, J.S., 1978. Independent dating in archaeological analysis, in: Schiffer, M.B. (Ed.), Advances in Archaeological Method and Theory. Academic Press, New York. volume 1, pp. 223–265.

Harary, F., Norman, R.Z., Cartwright, D., 1965. Structural Models: An Introduction to the Theory of Directed Graphs. John Wiley & Sons, New York.

Harris, E.C., 1989. Principles of Archaeological Stratigraphy. Second ed., Academic Press, London.

Harris, E.C., Brown, M.R., Brown, G.J. (Eds.), 1993. Practices of Archaeological Stratigraphy. Academic Press, London.

Herzog, I., 1993. Computer-aided Harris Matrix generation, in: Harris et al. (1993). chapter 13. pp. 201–217.

Herzog, I., Scollar, I., 1991. A new graph theoretic oriented program for Harris Matrix analysis, in: Lockyear, K., Rahtz, S. (Eds.), Computer Applications and Quantitative Methods in Archaeology: 1990. Tempus Reparatum, Oxford. number 565 in BAR International Series. chapter 9, pp. 53–59.

Hodder, I. (Ed.), 2007. Excavating Çatalhöyük: South, North and KOPAL Area Reports from the 1995–99 Seasons. Number 37 in British Institute at Ankara Monograph, McDonald Institute for Archaeological Research and British Institute at Ankara, Cambridge and London. Çatalhöyük Research Project Volume 3.

King, R., Morgan, B.J.T., Gimenez, O., Brooks, S.P., 2010. Bayesian Analysis for Population Ecology. Chapman & Hall/CRC, Boca Raton, FL.

Linick, T.W., Damon, P.E., Donahue, D.J., Jull, A.J.T., 1989. Accelerator mass spectrometry: The new revolution in radiocarbon dating. Quaternary International 1, 1–6.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS—a Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10, 325–337.

Orton, C., 1980. Mathematics in Archaeology. Collins, London.

Parent, E., Rivot, E., 2013. Introduction to Hierarchical Bayesian Modeling for Ecological Data. CRC Press, Boca Raton, FL.

Roskams, S., 2001. Excavation. Cambridge Manuals in Archaeology, Cambridge University Press, Cambridge.

Ryan, N.S., 1988. Data structures for stratigraphic analysis. Archaeological Computing Newsletter 14, 1–11.

Taylor, R.E., 1995. Radiocarbon dating: The continuing revolution. Evolutionary Anthropology 4, 169–181.

Towner, R.H., 2002. Archaeological dendrochronology in the southwestern United States. Evolutionary Anthropology 11, 68–84.

Triggs, J., 1993. The seriation of multilinear stratigraphic sequences, in: Harris et al. (1993). chapter 16. pp. 250–273.