



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/90295/>

Version: Accepted Version

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>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

- 1 **Title** Archaeological Sequence Diagrams and Bayesian Chronological Models
- 2 **Authors** Thomas S. Dye¹ and Caitlin E. Buck
- 3 **Affiliations** University of Hawai'i at Mānoa and University of Sheffield
- 4 **Address for correspondence** 735 Bishop St., Suite 315, Honolulu, HI 96813
- 5 **Email for correspondence** tsd@tsdye.com
- 6 **Approximate number of words** 5071

¹Corresponding author.

7 **Abstract**

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

20 **Keywords** Sequence diagram, Chronology, Directed graph, Bayesian radiocarbon calibra-
21 tion

22 Introduction

23 Advances in the methods and practice of radiocarbon dating in archaeology, sometimes
24 characterized as revolutionary (Bayliss, 2009; Taylor, 1995; Linick et al., 1989), have worked
25 generally to increase the precision of age estimates for archaeological events. A recent phase
26 of this radiocarbon revolution has as its focus Bayesian calibration (Buck et al., 1996), which
27 highlights the role of stratigraphic interpretation in the development of radiocarbon-based
28 site chronologies. A key innovation of Bayesian calibration is its ability to integrate ancillary
29 sources of chronological information with the information returned by the radiocarbon
30 dating laboratory. In a typical archaeological application having to do with site chronology,
31 records of the stratigraphic relationships of deposits and interfaces are a primary source of
32 this ancillary information. Common sense indicates that a site chronology based on "the
33 dates" *and* "the archaeology" is bound to be more reliable than one that relies only on one
34 or the other (Bayliss, 2009, 127). The improvement yielded by Bayesian calibration has been
35 demonstrated, perhaps most convincingly for the early Neolithic period of Southern Britain
36 and Ireland where time-scales with resolutions that approach a human generation have
37 been achieved (Bayliss et al., 2011). At Çatalhöyük, a Neolithic village in Anatolia, a basic
38 goal of the Bayesian calibration is to provide "calendar date estimates for the construction,
39 use, and disuse of the excavated buildings, in order to infer a structural narrative between
40 buildings that are not stratigraphically related" (Bayliss et al., 2014, 69). Given that a typical
41 house at Çatalhöyük was constructed, used, and disused over a period on the order of
42 60–145 years (Bayliss et al., 2014, 89), the ambitious goal of identifying contemporary houses
43 from spatially separate parts of the site without the aid of dendrochronology (Towner, 2002)
44 would have been wildly unrealistic prior to the development of AMS dating and Bayesian
45 calibration.

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

56 This paper describes a transformation algorithm based on the theory of directed graphs
57 that takes as its input a suitably structured sequence diagram and information on potential
58 dating samples to produce a chronological model for use in Bayesian calibration. To
59 demonstrate its utility in automating the creation of Bayesian chronological models, we
60 apply the algorithm to Buildings 1 and 5 in the North Area at Çatalhöyük (Cessford,

61 2007d,c,b,a). This example represents a relatively rare situation where a detailed sequence
62 diagram is published (Bayliss et al., 2014, Fig. 3.17) and dating specialists have carried out
63 several Bayesian calibrations (Cessford et al., 2005; Bayliss et al., 2014).

64 **Computing the Sequence Diagram**

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

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

83 Available Harris Matrix software packages are closed-source and do not permit program-
84 matic access to the DAG representation, so it proved necessary to develop the open-source
85 software package, *hm*, to achieve this goal (provided as supplementary material). Although
86 computer programmers quickly recognized that the sequence of observed stratigraphic
87 relationships at the heart of the sequence diagram can be represented as a DAG (Ryan,
88 1988; Herzog, 1993; Herzog and Scollar, 1991), the display conventions of the Harris Matrix
89 are tied to the layout of paper forms developed in the 1970s (Harris, 1989, 34) and these
90 conventions introduce complexities that can not be represented by a DAG. Thus, the *hm*
91 software abandons certain display conventions of the Harris Matrix in order to preserve a
92 pure DAG representation of the sequence diagram.

93 The following sections compare and contrast DAG and Harris Matrix representations of
94 the sequence diagram and present the data inputs to the *hm* software as tables that define
95 entities in a relational database (fig. 1). The first three sections consider the relationships
96 between contexts recognized by the Harris Matrix— i) no direct stratigraphic relationship, or
97 context identity, ii) an observed relationship of superposition, and iii) parts of a once-whole

98 context—in turn, as steps in the construction of a sequence diagram. This is followed by a
 99 consideration of periods and phases, which are conceptually similar interpretive constructs.

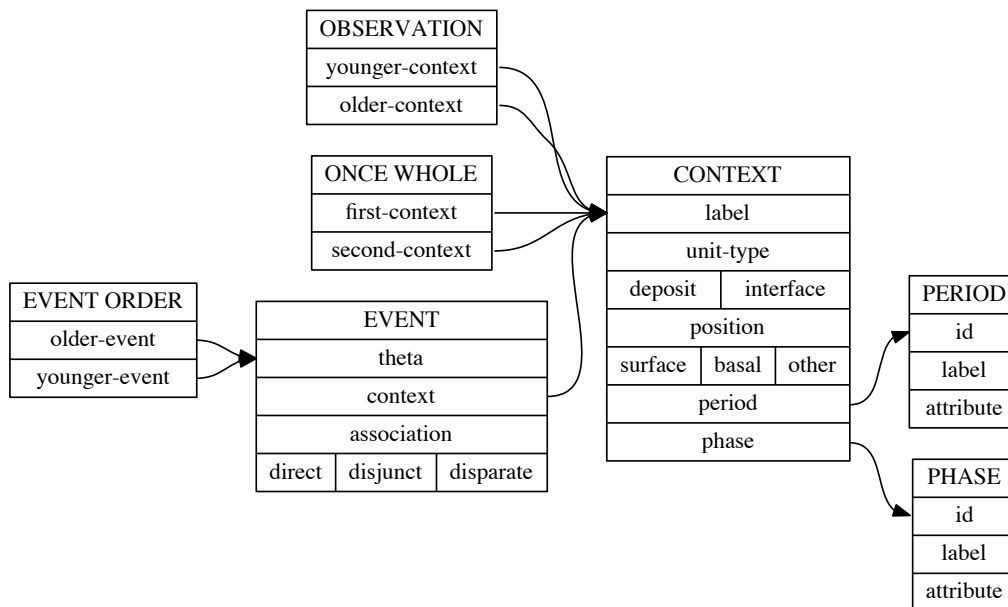


Figure 1: Relational database design for the seven tables of information used to construct stratigraphic and chronological directed graphs. Note that table names are uppercase, column names are lowercase, and divided entries define the domain of the column whose name is directly above, e.g., the `unit-type`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

100 Identification of Contexts

101 Archaeologists commonly identify five types of context: deposits, horizontal feature inter-
 102 faces, vertical feature interfaces, upstanding layer interfaces, and horizontal layer interfaces.
 103 The Harris Matrix was designed, in part, to ensure that all of the contexts identified at a
 104 site are included in the sequence (Roskams, 2001, 157) and to replace the previous archaeo-
 105 logical practice of recording contexts and their relationships with section drawings, which
 106 typically take in only some small fraction of the contexts identified at a site (Bibby, 1993,
 107 108).

108 In practice, the archaeologist working with a printed Harris Matrix sheet draws up a list

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

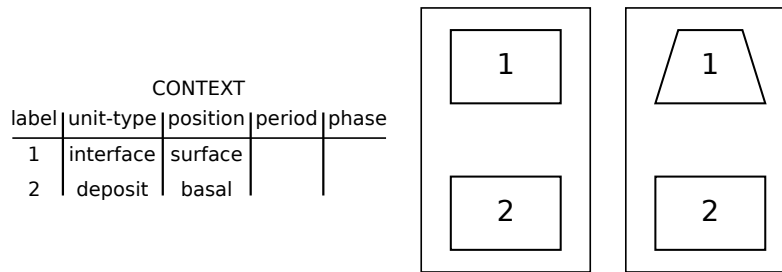


Figure 2: Initial stage in construction of a sequence diagram consisting of an interface, context 1, and a deposit, context 2: *left*, a five-column context table that records information about contexts (see fig. 1); *center*, Harris Matrix; *right*, directed graph.

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

133 Because the representation of a directed graph is not constrained by the conventions of
 134 the Harris Matrix, the shapes of nodes can express the fundamental distinction between
 135 depositional and interfacial contexts. The convention adopted here uses a rectangular box,

136 similar to the symbol used in a Harris Matrix, when `unit-type` is set to `deposit` and a
 137 trapezium when `unit-type` is set to `interface` (fig. 2, *right*).

138 Observed Stratigraphic Relationships

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

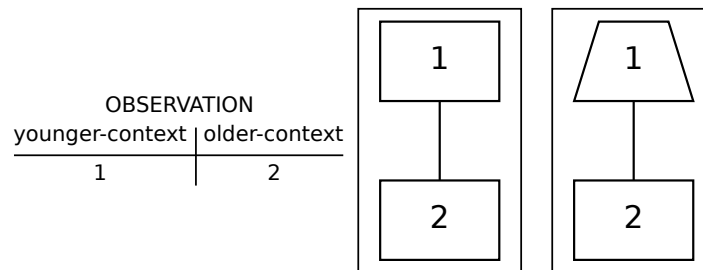


Figure 3: The sequence diagram after stratigraphic relationships are indicated with vertical lines: *left*, a two-column observation table that records the stratigraphic relationship between contexts 1 and 2 (see fig. 1); *center*, a Harris Matrix showing a younger interface, context 1, overlying an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

149 The directed graph uses the same table of observed stratigraphic relationships that the
 150 stratigrapher uses to draw the Harris Matrix. It is easy to see that each row of the table
 151 (fig. 3, *left*) represents an ordered pair of nodes, which in the theory of directed graphs
 152 defines an arc. The ordering is given by the stratigraphic relationship of the nodes; the
 153 younger context is by convention designated the start node of the arc and the older context
 154 the end node. It is customary to represent the arcs in a directed graph as arrows, with
 155 an arrowhead at the end of each arc to indicate direction. However, the Harris Matrix
 156 convention that uses a plain line and indicates direction by vertical position, such that
 157 a younger context appears above an older context with which it shares a stratigraphic

158 relationship, is appreciated by archaeologists who see in it the physical relationship of
159 the contexts when viewed in section. Thus, the directed graphs presented here adopt this
160 convention and draw arcs as lines rather than arrows (fig. 3, *right*).

161 At this stage in its construction, the Harris Matrix is a partial order, or poset (Orton, 1980,
162 67). The stratigraphic relationships that it records are irreflexive, because an archaeological
163 context cannot be stratigraphically superior or inferior to itself, asymmetrical because a
164 context that is stratigraphically superior to another context cannot be stratigraphically
165 inferior to it, and transitive because, given three contexts, 1, 2, and 3, if 1 is stratigraphically
166 superior to 2, and 2 is stratigraphically superior to 3, then 1 is stratigraphically superior to
167 3.

168 **Parts of Once-Whole Contexts**

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

183 **Stratigraphic Periods and Phases**

184 The terms “period” and “phase” are defined variously and sometimes interchangeably
185 by archaeologists. For the Harris Matrix, a “phase” groups contexts of similar age, and a
186 “period” groups phases of similar age, yielding a nested series of time intervals (Harris,
187 1989, 158). Defined in this way, both phases and periods are interpretive constructs that are
188 typically formulated with both stratigraphic and non-stratigraphic information. Because
189 “phase” is also used to describe Bayesian chronological models, here we use the term
190 “stratigraphic phase” to refer to a group of contexts, and the term “chronological phase” to
191 refer to a time period in a chronological model.

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

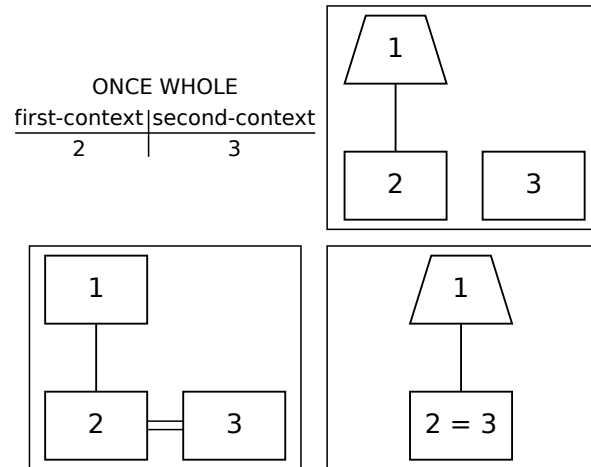


Figure 4: Three graphical representations of parts of a once-whole context: *top left*, two-column once-whole table recording the inference that contexts 2 and 3 are parts of a once-whole context (see fig. 1); *bottom left*, the Harris Matrix connects contexts 2 and 3 with two horizontal lines; *top 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”.

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

198 Structure of a Bayesian Chronological Model

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

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

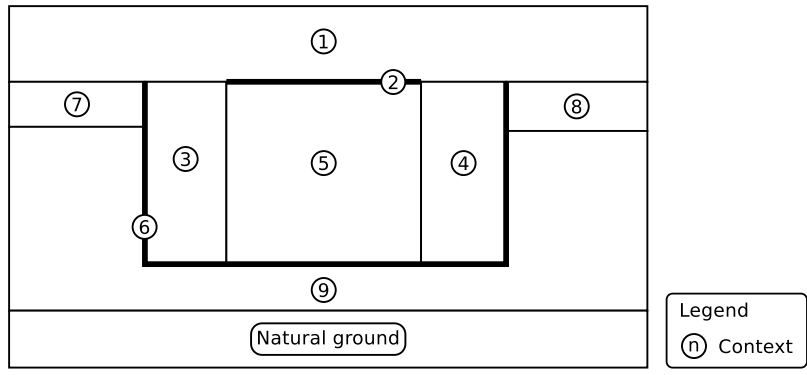


Figure 5: Illustrative stratigraphic profile drawing. Adapted from Harris (1989, Fig. 12a).

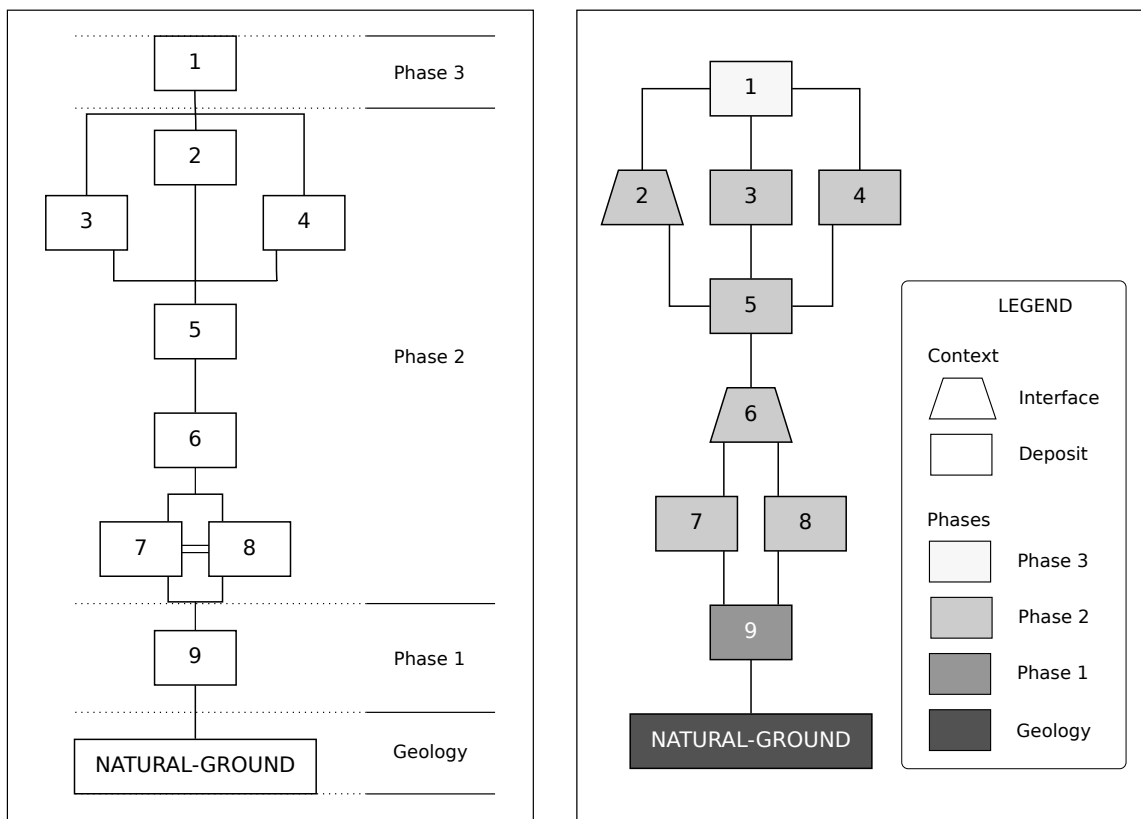


Figure 6: Hypothetical phasing of an example sequence developed by Harris (see fig. 5): *left*, the Harris Matrix representation, after Harris (1989, Fig. 12c); *right*, directed graph representation with nodes shaded to indicate phases.

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

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

- 217 • One chronological phase can be older than the other such that the end date for the
218 older chronological phase is older than the start date for the younger chronological
219 phase (fig. 7, *left*). This relationship, where a time interval separates two chronological
220 phases, arises in archaeological stratigraphy when two contexts are found on the same
221 line of a (possibly multi-linear) sequence but are separated by one or more contexts.
222 This relationship is relatively common in practical Bayesian chronological models.
223 Contexts that lack dating material are typically ignored in a Bayesian chronological
224 model.
- 225 • One chronological phase can be older than the other such that the end date for the
226 older chronological phase is the same age as the start date for the younger chrono-
227 logical phase (fig. 7, *middle*). This abutting relationship describes the relationship of
228 superposition that archaeologists typically observe in the field.
- 229 • One chronological phase can be older than the other such that the end date for the
230 older chronological phase is younger than the start date for the younger chronological
231 phase (fig. 7, *right*). This overlapping relationship cannot be determined solely on
232 stratigraphic grounds because the two contexts must be from different lines of a
233 multi-linear stratigraphic sequence. Other information, perhaps having to do with
234 the content of the contexts, is required to posit this kind of relationship (Triggs, 1993).

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

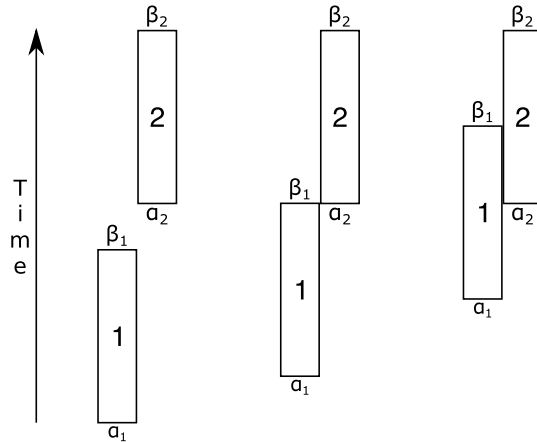


Figure 7: Schematic representation of three possible relationships between older and younger chronological phases: *left*, chronological phases 1 and 2 separated, $\beta_1 > \alpha_2$; *middle*, chronological phases 1 and 2 abutting, $\beta_1 = \alpha_2$; *right*, chronological phases 1 and 2 overlapping,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

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

248 Mapping a Sequence Diagram to a Chronological Model

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

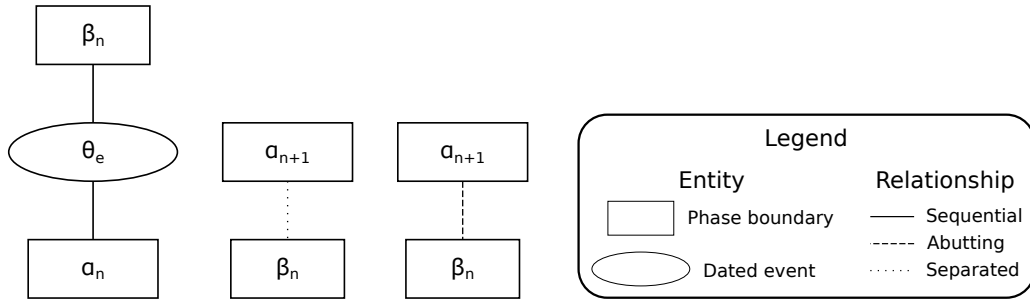


Figure 8: Entities and relationships of Bayesian chronological models represented as directed graphs: *left*, a chronological phase with a single dated event; *middle*, relationship between boundary parameters of separated chronological phases; *right*, relationship between boundary parameters of abutting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

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

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

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

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

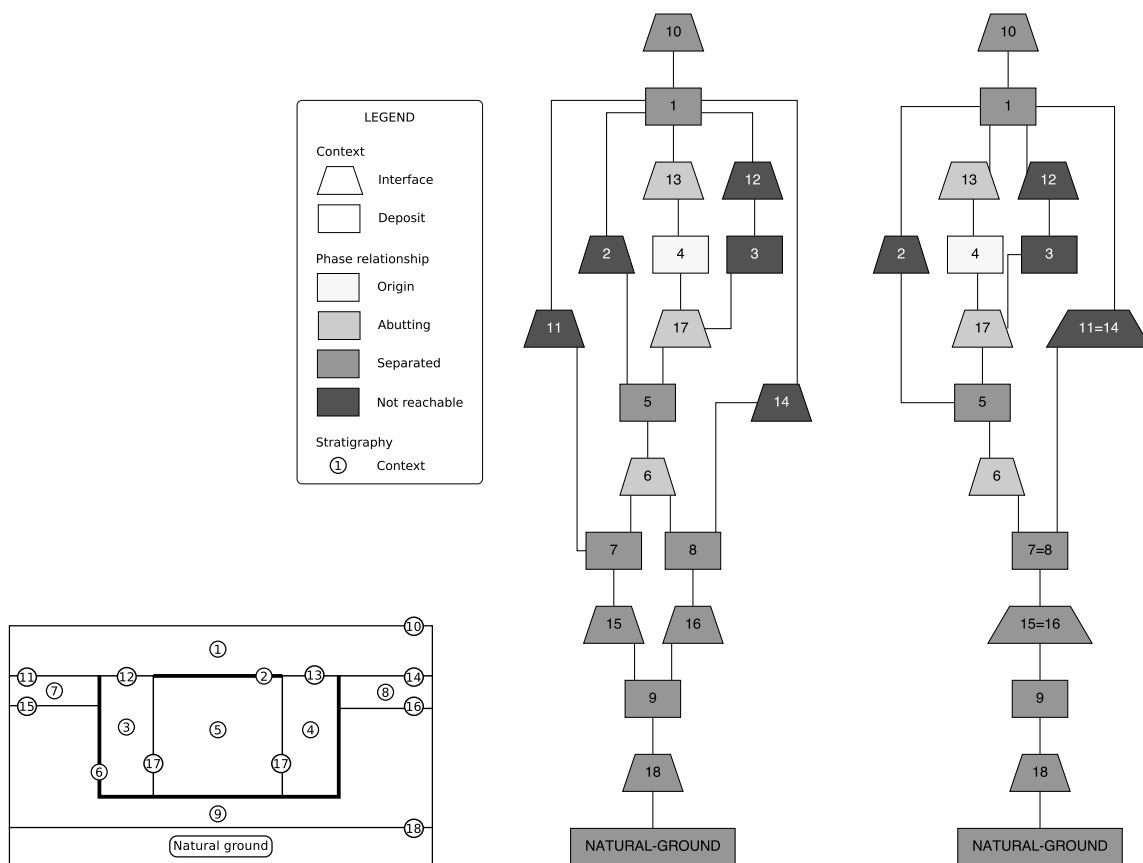


Figure 9: A hiatus-free sequence diagram with contexts shaded according to their chronological relationship to Context 4: *left*, the stratigraphic profile after Harris (1989, Fig. 12), with layer interfaces numbered 10–18 (cf. Fig. 5); *center*, a directed graph representation of the sequence diagram 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.

- 288 \mathcal{E} was excavated from a context and so is associated with one and only one member of
 289 $\mathcal{D} = \{d_1, d_2, \dots, d_D\}$, where $D = \#\mathcal{D}$.
- 290 1. For each member of \mathcal{D} , d_i , add two nodes to the chronological directed graph, one for
 291 the early boundary date, α_{d_i} , and the other for the late boundary date, β_{d_i} .
 - 292 2. For each member, e_j , of \mathcal{E} add one node, θ_{e_j} , to the chronological directed graph to
 293 represent its absolute date.
 - 294 3. For each row in the event order table add an arc from the younger node to the older
 295 node.
 - 296 4. For each row, $j = 1, 2, \dots, E$, of the event table (associated with archaeological context
 297 d_i and event with absolute date θ_{e_j}):
 - 298 a) if the indegree of θ_{e_j} is 0 (and association is not equal to “disjunct”) add an arc
 299 from α_{d_i} to θ_{e_j} and assign it a value of 0;
 - 300 b) if the outdegree of θ_{e_j} is 0 (and association is not equal to “disparate”) add an
 301 arc from θ_{e_j} to β_{d_i} and assign it a value of 0.
 - 302 5. For each pair (d_l, d_m) of archaeological contexts in the event table:
 - 303 a) if d_l is reachable from d_m in the directed graph of the archaeological sequence,
 304 add an arc from β_{d_l} to α_{d_m} in the chronological directed graph;
 - 305 b) if context d_m is adjacent to context d_l in the directed graph of the archaeological
 306 sequence, assign the arc from β_{d_l} to α_{d_m} in the chronological directed graph a
 307 value of 1, else assign it a value of 2;
 - 308 c) If context d_l is reachable from context d_m in the directed graph of the archae-
 309 ological sequence, add an arc from β_{d_m} to α_{d_l} in the chronological directed
 310 graph;
 - 311 d) if context d_l is adjacent to context d_m in the directed graph of the archaeological
 312 sequence, assign the arc from β_{d_m} to α_{d_l} in the chronological directed graph a
 313 value of 1, else assign it a value of 2.
 - 314 6. Perform transitive reduction.

315 Discussion

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

323 An example where the authors do discuss the impact of model choice is the work un-
324 dertaken to establish the chronology of Buildings 1 and 5 in the North Area excavations at
325 Çatalhöyük (Cessford et al., 2005; Bayliss et al., 2014). The initial work was exploratory in
326 nature, with one goal “to determine which types of material and/or context provide good
327 dating evidence” (Cessford et al., 2005, 84). The reliability of each dated sample was ranked
328 as “low” where “there is a direct stratigraphic relationship between determinations that
329 contradicts the relationship between the ages of the two determinations” (Cessford et al.,
330 2005, 76), “high” where the sample comes from “a consistently dated stratigraphic sequence”
331 (Cessford et al., 2005, 76) or where it is “short lived material from a context with a low prob-
332 ability of residuality” (Cessford et al., 2005, 76), or “medium” otherwise (Supplementary
333 Material Table S1). Where possible, contradictions were resolved with reference to four
334 of the five age determinations from Context 1332+² in Building 1, a “deliberately-placed
335 deposit of lentils which represents a single year’s harvest of a short-lived species that was
336 purposefully burnt” (Cessford et al., 2005, 86). Context 1332+ has a direct stratigraphic
337 relationship with all of the contexts excavated from Building 5, which underlies Building 1,
338 but its age relative to most of the contexts from Building 1 cannot be determined (fig. 10).

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

344 A directed graph representation of the chronological model implied by the exploratory
345 analysis accepts the assumption that each dated sample is associated with the context from
346 which it was collected (fig. 11). The chronological model indicates that none of the related
347 contexts superior to Context 1332+ in Building 1 were dated. Of the six dated contexts that
348 are stratigraphically related to Context 1332+, five are from Building 1 and one, Context
349 3810+, is from Building 5. Thus, potential contradictions could be worked out with direct
350 reference to the lentil deposit for a small subset of the dated contexts.

351 Carrying through the exploratory approach, Cessford et al. rejected the age determination
352 for one of the lentils, θ_{31} , as inconsistent with the other four age determinations on lentils
353 from Context 1332+, θ_{29} , θ_{30} , θ_{32} , and θ_{33} . Two dates on animal bone, θ_{42} from Context
354 1295a+ and θ_{24} from Context 1456, were assigned medium reliability because they were
355 older than botanical material from the same deposits and the four lentils (Cessford et al.,
356 2005, 88). As can be seen in Supplementary Material Figure S1, these comparisons with
357 the lentils are not based on stratigraphic relationships; Contexts 1295a+ and 1456 are not
358 reachable from Context 1332+ and their relative ages cannot be determined on stratigraphic

²It 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.

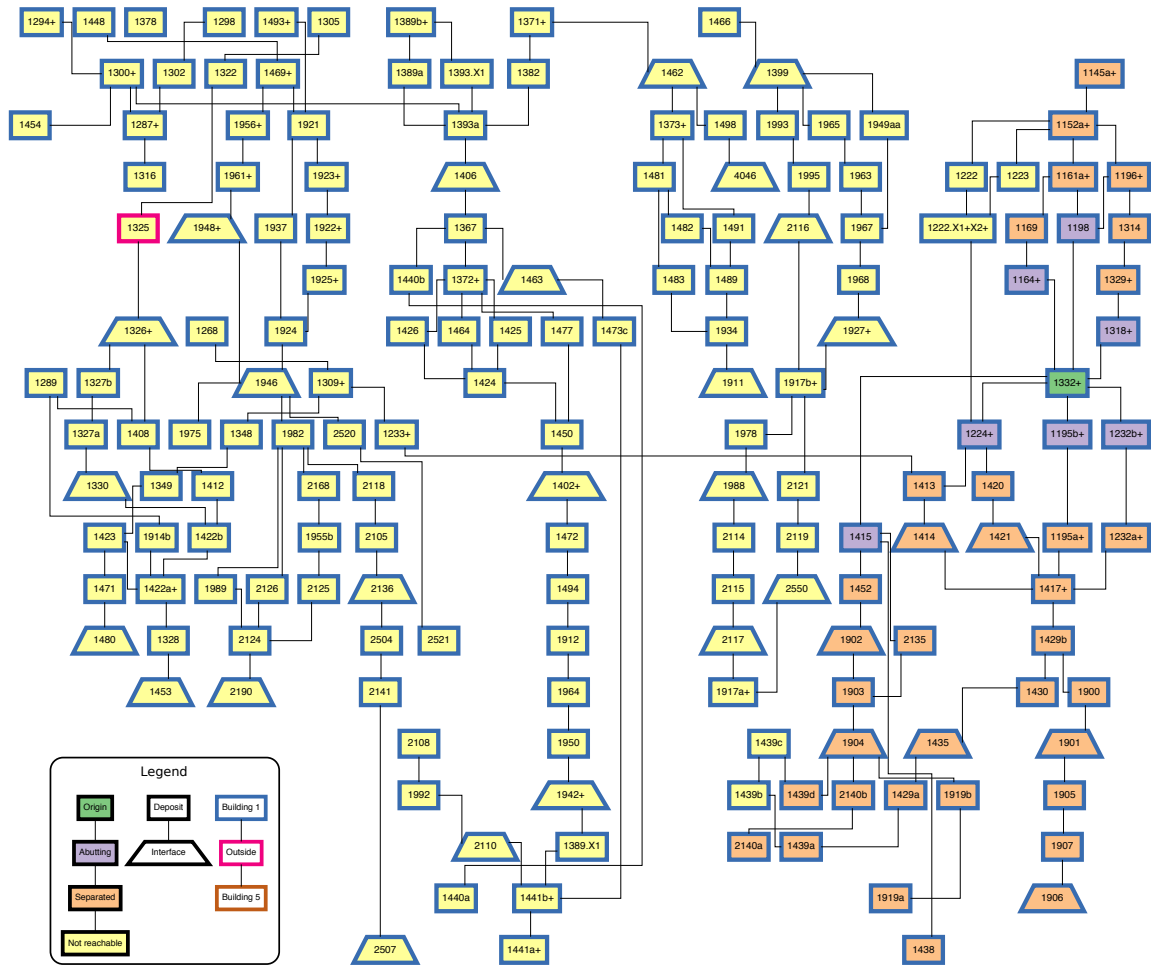


Figure 10: A portion of the sequence diagram for Buildings 1 and 5 of the North Area excavations at Çatalhöyük showing Context 1332+ in Building 1, adjacent and reachable contexts whose ages relative to Context 1332+ are known, and unreachable contexts whose ages relative to Context 1332+ can not be 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.

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

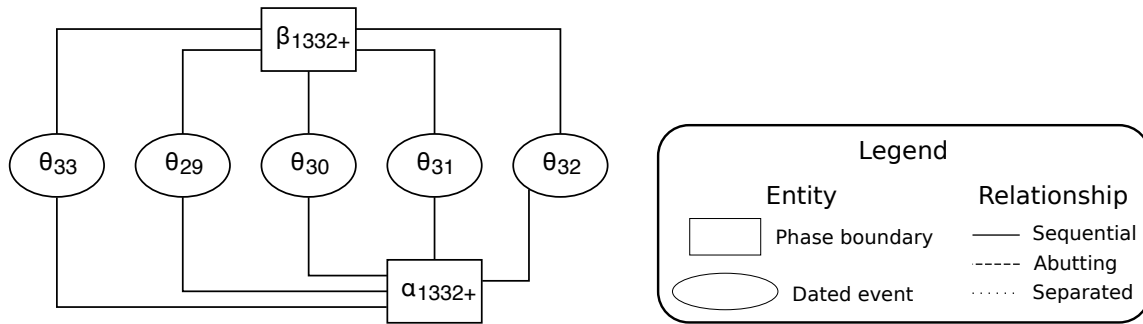


Figure 11: Representation of the dated lentils from Context 1332+ on a chronological model for determining which types of material and/or context provide good dating evidence using the dated samples reported by Cessford et al. (2005, Table 4.10) and the sequence diagram for the North Area excavations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

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

366 Subsequently, dates on human bone and antler processed at the Oxford Radiocarbon
 367 Accelerator Unit between 2000 and 2002 were shown to be incorrect due to a technical
 368 problem. When re-dated, the bone and antler samples from Çatalhöyük, including the
 369 six dates on human bone, were determined to be 50–150 BP younger than the original
 370 measurements (Bayliss et al., 2014, 79). In particular, θ_7 , which replaced θ_{48} from Context
 371 2529, stratigraphically inferior to the lentils, returned a date younger than the four lentils,
 372 but older than the lentil that was previously rejected. Accordingly, the four lentils previously
 373 determined to represent the true age of the lentil deposits were interpreted as residual, and
 374 the lentil previously believed to be a statistical outlier was accepted as dating the true age
 375 of the deposit. This circumstance, and a comprehensive reevaluation of the suitability of
 376 the dated sample materials based largely on experience gained subsequent to the original
 377 exploratory dating project (Bayliss et al., 2014, 81–88), resulted in a different chronological
 378 model, one in which a large proportion of the dated samples are *termini post quem* for the
 379 end date of the context from which they were collected but have no relationship to the start
 380 date (fig. 12). These “dangling θ ’s” graphically illustrate the substantial challenges posed
 381 by residuality for the ambitious dating project at Çatalhöyük.

382 Conclusions

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

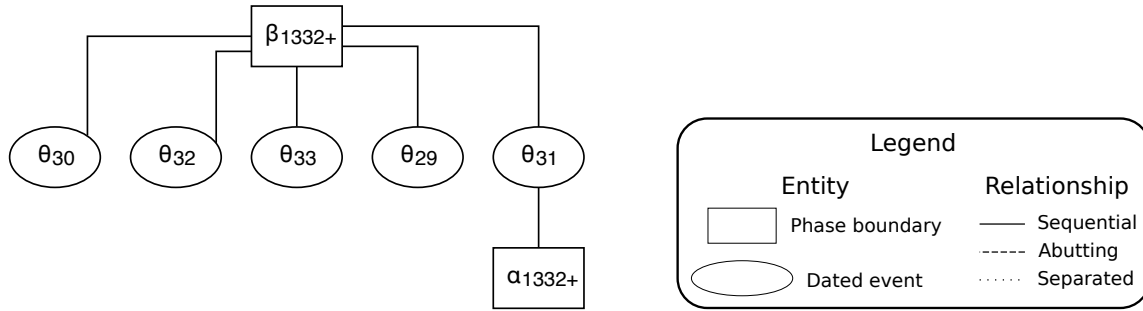


Figure 12: Representation of the dated lentils from Context 1332+ on a revised chronological model using the dated samples reported by Bayliss et al. (2014, Table 3.2) and the sequence diagram for the North Area excavations (fig. 10). The full revised chronological model is available as Supplementary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

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

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

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

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

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

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

413 **Acknowledgments**

414 The authors thank Alex Bayliss for suggesting the example of Çatalhöyük Buildings 1 and
415 5 and for providing guidance during our analysis; Craig Cessford for clarifying conven-
416 tions used in the representation of the Harris Matrix for Buildings 1 and 5 at Çatalhöyük;
417 Julian Richards, Keith May, and Kieron Niven for advice on data standards and assistance
418 searching the ADS archives; Eric Schulte for the `graph.lisp` library and for patient help
419 during development of the `hm` software; and two anonymous reviewers for recommending
420 that the paper include a real-world example. Any errors are the authors'.

421 References

- 422 Bayliss, A., 2009. Rolling out revolution: Using radiocarbon dating in archaeology. *Radiocarbon* 51, 123–147.
423
- 424 Bayliss, A., Farid, S., Higham, T., 2014. Time will tell: Practising Bayesian chronological
425 modeling on the East Mound, in: Hodder, I. (Ed.), *Çatalhöyük Excavations: The 2000–*
426 *2008 Seasons*. British Institute at Ankara and Cotsen Institute of Archaeology Press,
427 London and Los Angeles. number 46 in British Institute at Ankara Monograph, pp. 53–90.
428 *Çatalhöyük Research Project Volume 7*.
- 429 Bayliss, A., van der Plicht, J., Bronk Ramsey, C., McCormac, G., Healy, F., Whittle, A.,
430 2011. Towards generational time-scales: the quantitative interpretation of archaeological
431 chronologies, in: *Gathering Time: Dating the Early Neolithic Enclosures of Southern*
432 *Britain and Ireland*. Oxbow Books, Oxford. chapter 2, pp. 17–59.
- 433 Bibby, D.I., 1993. Building stratigraphic sequences on excavations: an example from
434 Konstanz, Germany, in: Harris et al. (1993). chapter 7. pp. 104–121.
- 435 Buck, C.E., Cavanagh, W.G., Litton, C.D., 1996. Bayesian Approach to Interpreting Archae-
436 ological Data. *Statistics in Practice*, John Wiley & Sons, Chichester, UK.
- 437 Carver, M., 1979. Three Saxo-Norman tenements in Durham City. *Medieval Archaeology*
438 23, 1–80.
- 439 Carver, M., 2005. Key ideas in excavation, in: Renfrew, C., Bahn, P. (Eds.), *Archaeology:*
440 *The Key Concepts*. Routledge, London, pp. 106–110.
- 441 Carver, M., 2009. *Archaeological Investigation*. Routledge, New York.
- 442 Cessford, C., 2007a. Building 1, in: Hodder (2007). pp. 405–530. *Çatalhöyük Research*
443 *Project Volume 3*.
- 444 Cessford, C., 2007b. Building 5, in: Hodder (2007). pp. 361–403. *Çatalhöyük Research*
445 *Project Volume 3*.
- 446 Cessford, C., 2007c. History of excavation of Buildings 1 and 5 and summary of phases, in:
447 Hodder (2007). pp. 345–360. *Çatalhöyük Research Project Volume 3*.
- 448 Cessford, C., 2007d. Overall discussion of Buildings 1 and 5, in: Hodder (2007). pp. 531–549.
449 *Çatalhöyük Research Project Volume 3*.
- 450 Cessford, C., Blumbach, M., Akoğlu, H.G., Higham, T., Kuniholm, P.I., Manning, S.W.,
451 Newton, M.W., Ozbakan, M., Ozer, A.M., 2005. Absolute dating at *Çatalhöyük*, in:
452 Hodder, I. (Ed.), *Inhabiting Çatalhöyük: Reports from the 1995–1999 Seasons*. McDonald

- 453 Institute for Archaeological Research and British Institute of Archaeology at Ankara,
454 Cambridge and London. number 38 in British Institute at Ankara Monograph, pp. 65–99.
455 Çatalhöyük Research Project Volume 4.
- 456 Clark, P., 2000. Negative features and interfaces, in: Roskams, S. (Ed.), *Interpreting Stratig-*
457 *raphy: Site Evaluation, Recording Procedures and Stratigraphic Analysis*. Archaeopress,
458 Oxford. number 910 in BAR International Series. chapter 11, pp. 103–105.
- 459 Dalland, M., 1984. A procedure for use in stratigraphic analysis. *Scottish Archaeological*
460 *Review* 3, 116–127.
- 461 Dean, J.S., 1978. Independent dating in archaeological analysis, in: Schiffer, M.B. (Ed.),
462 *Advances in Archaeological Method and Theory*. Academic Press, New York. volume 1,
463 pp. 223–265.
- 464 Harary, F., Norman, R.Z., Cartwright, D., 1965. *Structural Models: An Introduction to the*
465 *Theory of Directed Graphs*. John Wiley & Sons, New York.
- 466 Harris, E.C., 1989. *Principles of Archaeological Stratigraphy*. Second ed., Academic Press,
467 London.
- 468 Harris, E.C., Brown, M.R., Brown, G.J. (Eds.), 1993. *Practices of Archaeological Stratigraphy*.
469 Academic Press, London.
- 470 Herzog, I., 1993. Computer-aided Harris Matrix generation, in: Harris et al. (1993). chap-
471 ter 13. pp. 201–217.
- 472 Herzog, I., Scollar, I., 1991. A new graph theoretic oriented program for Harris Matrix anal-
473 ysis, in: Lockyear, K., Rahtz, S. (Eds.), *Computer Applications and Quantitative Methods*
474 *in Archaeology: 1990. Tempus Reparatum*, Oxford. number 565 in BAR International
475 Series. chapter 9, pp. 53–59.
- 476 Hodder, I. (Ed.), 2007. *Excavating Çatalhöyük: South, North and KOPAL Area Reports from*
477 *the 1995–99 Seasons*. Number 37 in British Institute at Ankara Monograph, McDonald
478 Institute for Archaeological Research and British Institute at Ankara, Cambridge and
479 London. Çatalhöyük Research Project Volume 3.
- 480 King, R., Morgan, B.J.T., Gimenez, O., Brooks, S.P., 2010. *Bayesian Analysis for Population*
481 *Ecology*. Chapman & Hall/CRC, Boca Raton, FL.
- 482 Linick, T.W., Damon, P.E., Donahue, D.J., Jull, A.J.T., 1989. Accelerator mass spectrometry:
483 The new revolution in radiocarbon dating. *Quaternary International* 1, 1–6.
- 484 Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS—a Bayesian modelling
485 framework: Concepts, structure, and extensibility. *Statistics and Computing* 10, 325–337.

- 486 Orton, C., 1980. *Mathematics in Archaeology*. Collins, London.
- 487 Parent, E., Rivot, E., 2013. *Introduction to Hierarchical Bayesian Modeling for Ecological*
488 *Data*. CRC Press, Boca Raton, FL.
- 489 Roskams, S., 2001. *Excavation*. Cambridge Manuals in Archaeology, Cambridge University
490 Press, Cambridge.
- 491 Ryan, N.S., 1988. Data structures for stratigraphic analysis. *Archaeological Computing*
492 *Newsletter* 14, 1–11.
- 493 Taylor, R.E., 1995. Radiocarbon dating: The continuing revolution. *Evolutionary Anthro-*
494 *pology* 4, 169–181.
- 495 Towner, R.H., 2002. Archaeological dendrochronology in the southwestern United States.
496 *Evolutionary Anthropology* 11, 68–84.
- 497 Triggs, J., 1993. The seriation of multilinear stratigraphic sequences, in: Harris et al. (1993).
498 chapter 16. pp. 250–273.