Journal of Archaeological Science 63 (2015) 84–93

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

Archaeological sequence diagrams and Bayesian chronological models

Thomas S. Dye [a], [*], Caitlin E. Buck [b]

[a] University of Hawai'i at Mānoa, 735 Bishop St., Suite 315, Honolulu, HI 96813, USA
[b] School of Mathematics and Statistics, University of Sheffield, UK

ARTICLE INFO

Article history:
Received 13 November 2014
Received in revised form
9 August 2015
Accepted 10 August 2015
Available online 12 August 2015

Keywords:
Sequence diagram
Chronology
Directed graph
Bayesian radiocarbon calibration

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.

1. 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

---

[*] Corresponding author.
E-mail address: tsd@tsdye.com (T.S. Dye).

http://dx.doi.org/10.1016/j.jas.2015.08.008
0305-4403/© 2015 Elsevier Ltd. All rights reserved.

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).

2. 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.

2.1. 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*).

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*).

2.2. 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*).

The directed graph uses the same table of observed stratigraphic