Deducing Event Chronology in a Cultural Heritage Documentation System

Jon Holmen[1] and Christian-Emil Ore[1]

[1] Department of Linguistics and Nordic Studies, University of Oslo, Oslo. Norway.

Abstract

In historically oriented research like archaeology, the determination of the chronology of events in the past plays an important role. For example, a fire of a house can seal off the layers physically below and give a partial relative dating of these. A well known tool in this area is the Harris Matrix, used to systematize the contexts and layers found in an excavation. In this paper we will discuss a related but more general tool for documenting and analyzing temporal entities like events. This tool is developed as a module of a four-dimensional event-oriented documentation database based on the conceptual model CIDOC-CRM (ISO21127). The database was developed for an archaeological excavation project in Western Norway. In addition to places, events, and actors, the database is designed to contain texts, images and maps used to document such entities. In use, the system will contain a dataset of events, their time-spans and relations between events. The system can detect conflicting dating, increase precision of beginnings, ends, and durations of events, and finally display a spatial and chronological overview. Given a time and a place within the dataset, the system can display all possible chronologies for the events in the set. So far, this tool has shown great potential, being used in projects involving large amounts of archival material as preparation for new excavations. Further development includes the possibility of using other temporal constraints, such as durations, and exploring the potential of adding spatial constraints and constraints on actors.

Keywords: *temporal reasoning, ontology, event*

1 INTRODUCTION

The objects on display in the exhibition halls are what most visitors associate with a museum. In reality, the exhibitions are only a fraction of the museum's collections and disseminate little of the information in the institution's archives and libraries. The old reports, catalogues, photos, and grey literature in general form the most important source of information about the museum collections and document the work of scholars in the field and at the museums. A diary tells part of the day to day story of an expedition. A report of an archaeological excavation will usually contain the archaeologist's interpretation of the site, told as the story based on the finds and structures excavated. In both cases, the basic structure of the texts is a series of events connecting objects, places and persons. Events, objects, places, and persons mentioned in one text or records may be referred to ("co-referenced") in other texts or records, where the stories continue. Thus other, more complete, stories emerge from complementary facts spread out over numerous primary sources. All cultural historical research starts with collecting these related pieces until the collected material covers the story the scholar is interested in revealing. Hence, core to scholarly work in general is this notion of "story," which is a way of putting things, people and events in a context of interaction, influence and reason.

An important part of the scholarly work is to establish the relative chronology of the events. The system discussed in this paper is meant to be an assistant for the working scholar to find possible chronologies. The system is based on the digitalization and documentation system built for the Norwegian University Museums as a part of the Norwegian Museum Project, 1998–2006.[1] The tool is a module of a four-dimensional event-oriented documentation database based on the conceptual model CIDOC-CRM, ISO21127.[2] The documentation database was originally developed for an archaeological project in Western Norway.

2 CIDOC-CRM

CIDOC-CRM is a formal ontology intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information. It was developed by interdisciplinary teams of experts, coming from fields such as computer science, archaeology, museum documentation, history of art, natural history, library science, physics, and philosophy, under the aegis of the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM).

[1]Jon Holmen et al., "From XML-tagged Acquisition Catalogues to an Event-based Relational Database," *Proceedings of Computer Applications in Archaeology 2004, Prato, Italy* (forthcoming). www.edd.uio.no/artiklar/arkeologi/jordal_caa2004.pdf; Christian-Emil Ore and Oddrun Rangsæter, *Final Report for the Museum Project 2007.* www.muspro.uio.no/engelsk-omM.shtml (seen 01-06-2009).

[2]Nicholas Crofts et al., eds. *Definition of the CIDOC Conceptual Reference Model, 2004–2009*. http://cidoc.ics.forth.gr/docs/cidoc_crm_version_5.0.1_Mar09.pdf (seen 05-27-2009).

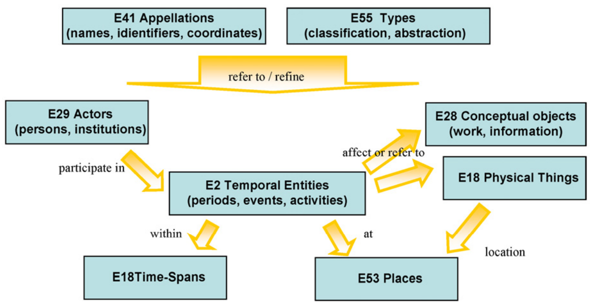

Figure 1. *The CIDOC-CRM, top classes for data integration.*

The CIDOC-CRM is event-centric core ontology in the sense that the model does not have classes for all particulars like, for example, the Getty's Art and Architecture Thesaurus with thousands of concepts. The central idea is that the notion of historical context can be abstracted as things, people and ideas meeting in space/time. The model contains, in addition, identification of real world items by real world names (appellations), a generalized classification mechanism (types), part decomposition of immaterial and physical things, temporal entities, groups of people (actors), places and time (time span), location of temporal entities in space-time and physical things in space, reference of information objects to any real world item (aboutness), and intellectual influence of things and events on human activities.

CIDOC-CRM is defined in an object-oriented formalism that allows for a compact definition with abstraction and generalization through the inheritance mechanisms (ISA hierarchy). CIDOC-CRM has 86 classes and 137 properties. The most central classes and properties for data interchange are shown in figure 1.

As an illustration of how events can be modeled in CRM, a traditional English wedding is used as an example. As shown in figures 1 and 2, the event is at the core of the model, both conceptually and visually. Through this, the event occurring at a specific location, the persons, the groom, the bride and the groom's best man are connected. Their roles and the event itself are classified by the types (e.g. selected from a thesaurus).
In an information system, the formal relation between the spouses has to be deduced by checking whether the two have participated in a wedding event in the role of bride and groom. Alternatively, one may introduce short-cut relations indicated in grey at the bottom.

3 The temporal analysis and the CIDOC-CRM

In the CIDOC-CRM there are basically two ways to express chronological information about events. They can be dated relatively to each other: before, after, overlapping, etc. (CRM properties P114, P117-120). The events are seen as wholes with a clear-cut beginning and end.

The events can also be connected to a timeline through time-span(s); see figure 3. An E51 Time-Span represents a temporal extension on the abstraction level, while an E61 Time Primitive is on the implementation level and is typically represented as an interval on a timeline. An event has in principle only one time-span. If an event has more than one time-span, this expresses divergent opinions about the temporal extension of the event. Conversely, two events may have a common time-span. This expresses the fact that the events occurred simultaneously.

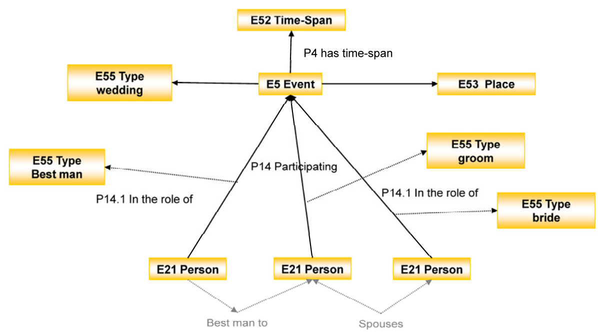

Figure 2. *A CRM diagram for a traditional English wedding.*

In cultural heritage databases, it is common to store time determinations as text, e.g. "the summer 1349" or "active 1450-1455". Such textually defined dates should be mapped to E51 Time-Span in the CIDOC-CRM. To enable queries, a widespread solution is to represent such text dates as intervals with a numerical start date and an end date. That is, the summer 1349 can be given the interpretation 15-04-1349 to 15-10-1349. This can of course be extended to all dates. The year 1349 is interpreted as the interval 01-01-1349 to 31-01-1349. The degree of accuracy can be extended *ad libitum*; the day 14-07-1789 is represented as 0000-14-07-1789 to 2400-14-07-1789.[1] The interpretation of dates as intervals corresponds to the E61 Time Primitive in the CIDOC-CRM. The interpretation of dates as intervals or time-spans is a powerful model and is easy to implement.

The CIDOC-CRM has a dual view on events. On the one hand they are viewed as entities that can be put in relation to each other by the time operators *P114 is equal in time to, P115 finishes, P116 starts, P117 occurs during, P118 overlaps in time with, P119 meets*

[1]Martin Doerr and Anthi Yiortsou, "Implementing a Temporal Datatype." *Technical Report ICS-FORTH/TR-236*, 1998. url: www.ics.forth.gr/isl/ publications/paperlink/ implementing_a _temporal_datatype.ps.gz.

in time with, P120 occurs before.[1] When using these operators, an event A can be thought to have an associated time-span with clear-cut start point A_s and end point A_e.

On the other hand, in scientific work a time-span represents an abstract approximation to the temporal extent of an event. The exact start and end points are usually not known. In the CIDOC-CRM this approximation to the temporal extension of events is modeled as two intervals on the timeline, one for the outer bounds (P81 *at sometime within*) and one for the inner bounds (P82 *ongoing throughout*); see figure 3. In an ideal situation with absolute accuracy, these two intervals should be identical. In most actual research they are not. In many cases there is only information about an outer interval (*at sometime within*), that is, by some terminus post quem (TPQ) and terminus ante quem (TAQ) for the event.

The properties *P81 at sometime within* and *P82 ongoing throughout* allow for an interpretation with fuzzy start and end points of time-spans. The start points of the two intervals define a third interval (A_{ss}, A_{se}) in which the event must have started. This interval can be interpreted as an approximation to the unknown exact start point of the event. Correspondingly, the end points of the inner and outer intervals define a fourth interval (A_{es}, A_{ee}) approximating the exact end point of the time-span for the event (fig. 3).

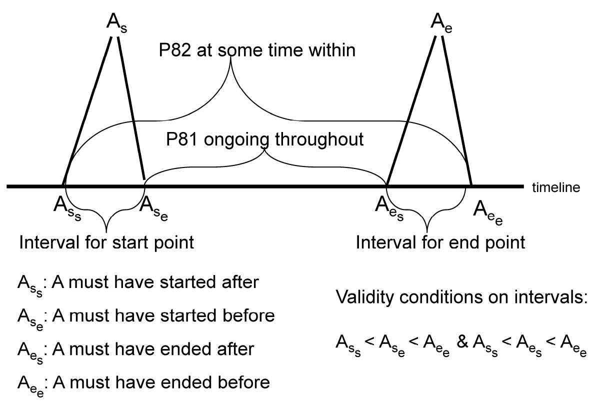

Figure 3. *A time-span for an event modeled as four points on the timeline.*

The four values or points on the timeline can be used to implement a reasoning system for time-spans, as well as deduction rules corresponding to the so-called Allen operators modeled as P114 to P120 in the CIDOC-CRM.

The objective of any deduction system is to increase the information by applying rules of interference. In our system, the objective is to deduce the possible relative chronology of events, as well as deducing the most accurate approximation possible to the actual dates for the events.

4 FUZZY POINTS IN TIME

The basic and most important assumption in the system described in this paper, is that a point in time only can be given as an approximation to, or fuzzy measurement of, the point and should be expressed as a time interval.

Our representation of a point in time, A_1, is expressed as $A_1 = (A_{1s}, A_{1e})$ where A_{1s} and A_{1e} are values on the timeline such that $A_{1s} < A_1 < A_{1e}$. A_{1s} is a TPQ, that is, the *earliest possible occurrence* of A_1, A_{1e} is a TAQ, that is, the *latest possible occurrence* of A_1:

$A' < A_{1s}$	=>	A' must be before A_1
$A_{1e} < A'$	=>	A' must be after A_1
$A_{1s} < A' < A_{1e}$	=>	A' can be before, equal to or after A_1

For a given point in time, e.g. A_s in figure 3, we increase the information value or strengthen our knowledge about A_s by shortening the interval (A_{ss}, A_{se}). By making the interval longer we will weaken the statement. Since this will decrease our knowledge, the deduction rules will never result in a lengthening of the interval (A_{ss}, A_{se}).

The deduction rules for time-spans are based on the so-called Allen operators between events. As we will discuss below, the Allen operators can be expressed as relations (before, equal, after) between the start and end points of the time-spans for the events (see figure 5). In the model, the start and end points of time-spans are seen as approximations and expressed as intervals. Correspondingly, the relations between start and end points are modeled as relations between intervals on the timeline.

Figure 4 shows how the intervals for two points in time, A_1 and A_2, can be adjusted to A'_1 and A'_2 according to the knowledge that $A_1 < A_2$. Note that the new intervals are always contained in the original ones and represent a better approximation. In the first case in figure 4, the intervals are disjoint and in the correct order. Here no new information is introduced by the fact $A_1 < A_2$. In the last case, the intervals are disjoint but in the wrong order and the introduction of the fact $A_1 < A_2$ leads to a contradiction. The inference rule $A'_1 = (A_{1s}, \min(A_{1s}, A_{2s}))$, $A'_2 = (\max(A_{1s}, A_{2s}), A_2)$ will result in $A'_{1e} < A'_{1s}$ which is meaningless in a model where a point is modeled as an interval approximating the real point in time.

[1] James F. Allen, "Towards a General Theory of Action and Time," *Artificial Intelligence* 23 (1984): 123–154; Nicholas Crofts et al., eds. *Definition of the CIDOC Conceptual Reference Model, 2004–2009*. http://cidoc.ics.forth.gr/docs/cidoc_crm_version_5.0.1_Mar09.pdf (accessed 05-27-2009).

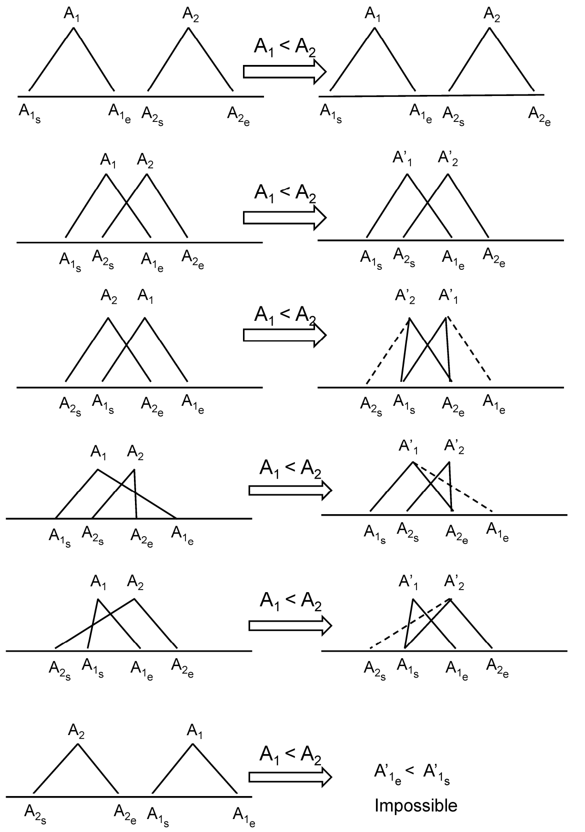

Figure 4. *Deduction rule for $A_1 < A_2$, where A_1 A_2 are two points in time modeled as intervals.*

5 TIME-SPANS WITH FUZZY BOUNDS

Figure 3 shows how a time-span A for an event is expressed in the approximation model. The start point A_s and end point A_e are expressed as the intervals (A_{ss}, A_{se}), and (A_{es}, A_{ee}). We require that the start point occurs before the end point, that is, $A_s < A_e$. As shown in figure 4, the only situation we do not allow is where the intervals are disjoint and (A_{es}, A_{ee}) is to the left of (A_{ss}, A_{se}). The diagrams in the right column of figure 4 show the possible constellations of the two intervals for the start and end point. These can be summarized as the validity requirement defined below.

Definition

A quadruple A = (A_{ss}, A_{se}, A_{es}, A_{ee}) is a valid implementation of a time-span, written valid(A) if

$$(A_{ss} \leq A_{se} \leq A_{ee}) \ \& \ (A_{ss} \leq A_{es} \leq A_{ee})$$

The validity requirement is compliant with the fact that the events extension on the timeline is limited by TPQ and TAQ. The start and end points A_{ss} and A_{ee} represent TPQ and TAQ for the event. In a deduction system, the information is weakened if the A_{ss} is moved to the left or the A_{ee} is moved to the right, since the period in which the event may have occurred is enlarged. Correspondingly the statement is weakened if the A_{se} is moved to the right or A_{es} is moved to the left since in this case the period in which the event must have been ongoing is shortened.

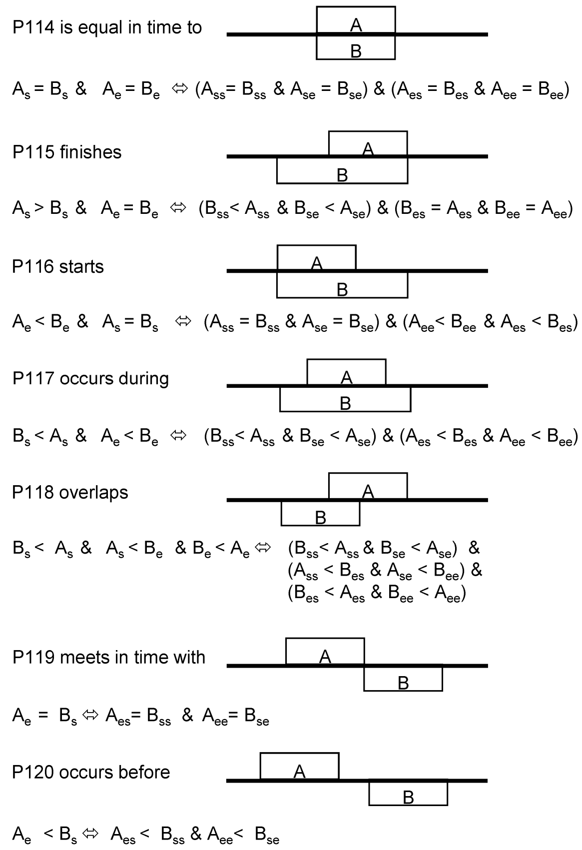

Figure 5. *Allen operators expressed as point to point relations.*

There is, however, no requirement that A_{se} should be less than A_{es}. The point A_{se} indicates only the knowledge that the event must have started *before* this point in time and the point A_{es} indicates the knowledge that the event must have ended *after* this point in time. If $A_{se} > A_{es}$, that is the length of the *P81 ongoing throughout* interval is negative, this only tells us that our information is a very incomplete approximation.

Figure 5 shows an overview of the CRM properties for the temporal relations between two events, their start and end points equivalents and the required adjustments for the fuzzy boundaries to implement the relations. The representation of dates as two intervals makes the implementation more complex, since each date has to be represented as four values on the time line. It is, however, possible to implement this model in a non-complex and efficient way.

Definition

Let E and F be events, A, A', B, B' be time-spans and (A_{ss},A_{se},A_{es},A_{ee}) etc. be the time-spans expressed as four points on the timeline.

If E has time-span A, F has time-span B, E occurs during F then E has time-span A’, F has time-span B’ where

A'_{ss} is the smallest value $\geq A_{ss}$,
A'_{se} is the largest value $\leq A_{se}$,
A'_{es} is the smallest value $\geq A_{es}$,
A'_{ee} is the largest value $\leq A_{ee}$
B'_{ss} is the smallest value $\geq B_{ss}$,
B'_{se} is the largest value $\leq B_{se}$
B'_{es} is the smallest value $\geq B_{es}$,
B'_{ee} is the largest value $\leq B_{ee}$

such that

$(A'_{ss} > B'_{ss} \,\&\, A'_{se} > B'_{se})$ &
$(A'_{es} < B'_{es} \,\&\, A'_{ee} < B'_{ee})$ &
valid(A’) & valid(B’)

The deduction rules for the remaining properties P114, P115, P116, P118, P119, and P120 are defined correspondingly.

Rule 1

Intersection of time-spans for the same event:

$$A \,\&\, B = (A_{ss}, A_{se}, A_{es}, A_{ee}) \,\&\, (B_{ss}, B_{se}, B_{es}, B_{ee}) = (\max(A_{ss},B_{ss}), \min(A_{se},B_{se}), \max(A_{es},B_{es}), \min(A_{ee},B_{ee}))$$

Constraint 1

$$\text{minlength}(A) = x \Rightarrow (\text{distance}(A_{se},A_{ee}) \geq x) \,\&\, (\text{distance}(A_{ss},A_{es}) \geq x)$$

Constraint 2

$$\text{maxlength}(A) = x \Rightarrow (\text{distance}(A_{se},A_{ee}) \leq x) \,\&\, (\text{distance}(A_{ss},A_{es}) \leq x)$$

The rule 1 is associative and together with the undefined time-span $(-\infty,-\infty,\infty,\infty)$ forms a monoid.

In other temporal systems, like the one defined by Cowley and Plexousakis,[1] the inner interval does not exist or is undefined in the cases where the length is negative. Our model is consistent with this solution, but preserves all information from the sources.

[1] W. Cowley and D. Plexousakis, “An Interval Algebra for Indeterminate Time,” in *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence* (Cambridge: MIT Press, 2000) 470–475.

6 EXAMPLE

The following example (see fig. 6) is not real, but indicates the main features of the algebra. Assume a document dated 1660 contains the minutes of a meeting where one approved the plans to build a church at Hillsend. Assume also that other documents state that Mr. and Mrs. Brown were married 24-06-1690 in the new church at Hillsend. If these two documents can be trusted, we may conclude that the building of the church at Hillsend took place some time between 1660 and 1690. This set is A_{ss} to 1660 and A_{es} to 1690. What about the inner interval? We clearly don’t know when the building actually was constructed, but it has to be some time inside the outer interval. It cannot have been finished before 1660 nor started after 1690. Therefore the system sets the earliest possible occurrence of the end (A_{es}) to 1660, and the latest possible occurrence of the start (A_{se}) to 1690. This may seem counter-intuitive. One should remember that the end points of the time-span are only intervals approximating the real value. When the *P81 ongoing throughout* interval is negative, this only tells us that our information is very incomplete. When more information is added the interval will become shorter and eventually get a positive length.

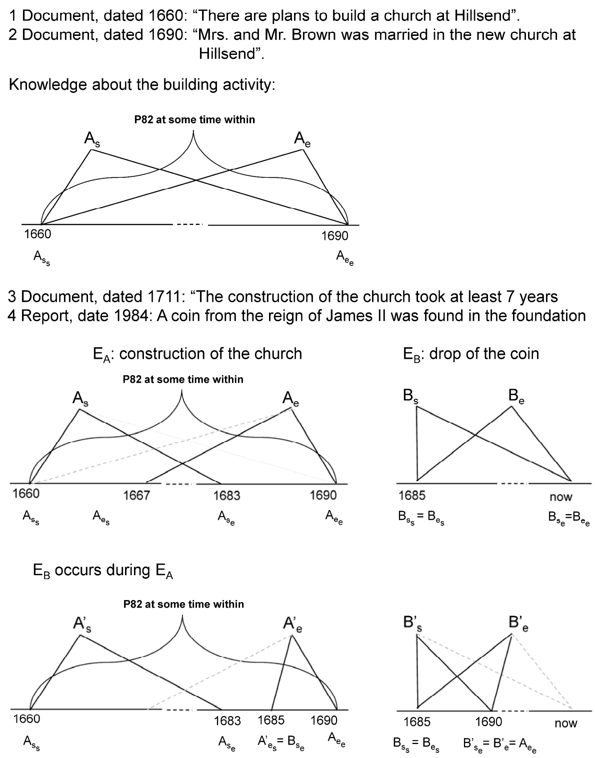

Figure 6. *Construction of a church.*

Now, assume we also have a document reporting about the building process stating that the construction of the church lasted seven years. When this information is added, then we know that the construction could have ended in 1667 at the earliest (A_{es}) and started in 1683 at the latest (A_{se}). Thus this extra fact adds precision although the length of the inner interval still is negative.

Finally, assume that during an excavation of the church in 1980 a coin from the reign of James II of England was found inside the foundation. Now we have several facts. James the II was King of England from 1685 to 1701. The minting of the coin took place between 1685 and 1701. The disposal of the coin can in principle have taken place from 1685 to the present. The fact that the coin was found sealed inside the foundation indicates that the disposal happened before the completion of the church. Thus the construction of the church was ongoing some time after 1685. This moves the earliest possible end point of the construction, A_{es}, to 1685 and the inner interval (ongoing throughout) is from 1683 to 1685.

The reasoning in the above example can be formalized by the formulation of a set of interference rules based on the Allen relations expressed in the CIDOC-CRM and the four-value model for time spans. The table in figure 5 expresses how Allen relations between events correspond to numerical relations between the four values for the corresponding time spans. On the basis of this correspondence we can formulate a series of applications of the deduction rules as follows:

1. A_1 time-span E_1 for the construction deduced from the minutes from meeting
   A, = (1660, ∞, 1660, ∞)
2. A_2 time-span E_1 for the construction deduced from the wedding document
   A_2 = (-∞, 1690, -∞, 1690)
3. A_3 time-span E_1 concluded from A& B by rule 1.
   A_3 = (1660, 1690, 1660, 1690)
4. A_4 time-span for concluded from A_3 using rule 2 and based on the document stating the duration was 7 years
   A_4 = (1660, 1683, 1667, 1690)
5. B time-span of E_2 the reign of James II,
   B = (1685,1685, 1701, 1701)
6. C time-span for E_3 the mining of the coin
7. E_3 occurs during E_2 gives
   C = (1685, 1701, 1685, 1701)
8. D time-span for the E_4 disposal of the coin
9. E_3 occurs before E_4, gives
   D = (1685, today, 1685, today)
10. E_3 occurs during E_1, gives
    A_5 = (1660, 1683, 1685, 1690)

7 THE DOCUMENTATION SYSTEM

The mechanism for time reasoning is a module in a larger documentation system based on CIDOC-CRM. The documentation system consists of several modules:

The document module. This part of the system is a documentation repository designed to store documents, measurements, maps, photos and so on. This unit is in itself event based. The events express the provenience of the documentation and are on a meta level. In our example, the documentation module would have kept facsimiles or transcriptions of (parts of) the original documents, for example the minutes dated 1660 of the meeting about building of the church, the documentation of the marriage in the church from 1690 and the find report from 1980 about the coin.

The event module. With this module the user can enter information about events and their dates and connected persons and places. The current system does not contain information about artifacts and other physical things. A posthole or a find is represented by the events, e.g., the digging of the posthole, backfilling and so on. This is similar to setting up a Harris matrix. In such a matrix the elements are the surfaces, and the purpose of the matrix is to establish the relative chronology of the events in which the surfaces were created. The underlying CIDOC-CRM compliant database contains structures for storing information about things. Thus it is possible to extend the current system such that it is possible to import data from an excavation database like the Swedish system INTRASIS.

All events in this repository are currently entered manually one by one. It is obligatory to give each fact a source of information, which in many cases will be a document in the system's repository, where the event is (implicitly) described. The detection of the events is based on a researcher's interpretation of the texts, and the system requires information about who did the extraction. This is necessary to facilitate a later reliability check and ensure the reproducibility of the results. There is, however, no restriction on the facts added. That is, the system may contain conflicting facts. This is intentional, since one of the purposes of the system is to help the users find conflicts in their source material.

The temporal analyzer module. This module helps the user to establish chronological order of the events represented by the facts in the system. The user can run the temporal analyzer on a set of events or add the facts one by one to the set used by the analyzer. The user adds temporal restrictions (P114-P120); for example, the disposal of coin *occurred within* construction of church. This is done through an interactive process. The temporal analyzer performs the point to point operations on the fuzzy endpoints of the time-spans as described above to accomplish the desired relations. Any kind of inconsistency between the attempted chronological order and the initial values in the time-span's primitives will be detected. The user is given the choice either to change the attempted chronological order or go back and review the dating entered in the event module in order to find the source of the contradicting information.

Stored Story Objects (SSO). The result of the interactive play in terms of added temporal relations between the events and resulting decreased fuzziness in the dating is kept separated from the data in the event repository. The latter remains unchanged.

It is important to note the difference between the two types of information. In our example, the time-span for the building of the church (A_4) is based on reading of documents in the document repository about the event. The final time-span deduced by the use of the system is the result of relating this event to other events based on other documents: the archaeological report, the classification of the coin, and the system's deduction rules. Any change in the time-span for the coin disposal might also result in a change for the dating of the construction of the church. Similarly, a change in classification of the coin, from James II to Charles II, might change the possible time-span for coin disposal, which again might change the time-span for the actual building period of the church.

The network of temporal related events and the resulting time-spans is stored as user defined "story objects". A stored story object (SSO) is an XML document containing the events (identifiers) and at least their relation and dates as adjusted by the process described above. In addition there is a unique reference to the original events in the event repository where the original time-span boundaries, actors, place and descriptions are kept.

In many cases there is more than one possible chronology. The use of SSOs allows the user to store an unlimited number of possible versions, all compliant with the original, documented facts. Different SSOs based on the same source of data reflect different opinions and should be used to analyze which facts should be further studied to possibly resolve the conflicting conclusions.

The overall state analyzer. The tool has a simple graphical interface for working with the time-spans. Through this interface it is possible also to invoke the overall state analyzer. The state analyzer parses through the timeline and splits it into segments where the overall possibilities of combinations of *before, during* and *after* between the time-spans remain unchanged through the whole segment. For each segment of time, the state analyzer is capable of finding all possible combined states at a given place. This is very useful when there are complex relations between time-spans in the same segment of time. The algorithm is based on analyzing each relation between points occurring in the time segment. All legal combinations of states between two points are then joined together to compute all possible overall states in the given segment of time.

7 POSSIBLE EXTENSIONS

The current system has seven deduction rules for reasoning about the relative chronology of two events based on the Allen operators (see fig. 5). In addition, there is a single rule for intersection of two time-spans for the same event and two deduction rules based on constraints on maximum and minimum duration. These rules are based on the core ontology expressed by the CIDOC-CRM.

When applied to more specific fields, it can be useful to sharpen the temporal reasoning by adding more domain-specific rules. However, a surprising amount of temporal rules and constraints can be expressed as a set of events with a chronological (partial) ordering based on the existing rules. The quite obvious requirement that a person must be allowed the necessary time to travel when participating physically in two events, E_1 and E_2, at separate places, can be expressed by introducing a third event, E_3, for the travel. The event E_3 must have a minimum duration depending on the travel distance between the two places at the given time period of history. In a future extension we will add a module in which the users can add such additional rules and constraints to the analyzer.

Another feature to explore is adding probability into the model. In its basic form the model operates only with absolutes. In the area of fuzziness, the possible states of *before* or *after* have equal weight. In more advanced models one might put a parameter of probability into these areas, e.g. a Gaussian or linear curve expressing the probability from first to last possible occurrence of a given point.

The temporal analyzer will be given an interface making it possible to integrate with other tools, such as excavation applications delivering Harris Matrix or Grey Literature repositories with tagged events. Any application capable of delivering data on our SSO format should be able to make use of the system. At the University of Tours, France, a system for documenting preindustrial Tours is being developed. This system is based on the so called OH_FET Model (Social Use, Space, and Time). The motivation behind our system and theirs is overlapping, although the angle of attack seems to be different. However, in our opinion it would be very beneficial to compare the two systems more systematically and establish a common set of tools for documenting the past.

8 SUMMARY

The model for time-span, based on CIDOC-CRM, handles fuzzy dating. The point to point algorithms for implementing temporal relations result in reduced fuzziness in the time-spans. The algorithms also detect any temporal inconsistency in the provided interpretation of the material. The possibility of including total undetermined time-spans with negative "ongoing throughout" makes any event a candidate for the system. The Stored Story Object (SSO) format

[1]Bastien Lefebvre et al., "Understanding Urban Fabric with the OH_FET Model Based on Social Use, Space and Time," *Archeologia e Calcolatori* 19 (2008): 195–214.

facilitates new possibilities for storing and fetching temporal networks of events and makes the system very flexible and multi-versioned. The tool shows great potential in integrating data from various sources into an event-based system and in analyzing and deducing event chronology.

ACKNOWLEDGEMENTS

We thank our colleagues, especially Øyvind Eide, at the Unit for Digital Documentation and in the Museum and Documentation project for the many interesting discussions about data integration and text encoding over the years. The working group meeting in the CRM SIG and the many valued colleagues in the museums have also been important sources of inspiration.

BIBLIOGRAPHY

Allen, James F. "Towards a General Theory of Action and Time," *Artificial Intelligence* 23 (1984): 123–154.

Cowley, W., and D. Plexousakis. "An Interval Algebra for Indeterminate Time," in *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence* (Cambridge: MIT Press, 2000) 470–475.

Crofts, Nicholas, Martin Doerr, Tony Gill, Stephan Stead, and Matthew Stiff (eds.) (2004–2009): *Definition of the CIDOC Conceptual Reference Model, 2004–2009*. http://cidoc.ics.forth.gr/docs/cidoc_crm_version_5.0.1 _Mar09.pdf (accessed May 27, 2009).

Doerr, Martin, and Anthi Yiortsou. "Implementing a Temporal Datatype." *Technical Report ICS-FORTH/TR-236*. 1998. url: www.ics.forth.gr/isl/publications/paperlink/implementing_a_temporal_datatype.ps.gz (accessed May 27, 2009).

Holmen, Jon, Ellen Jordal, Stein Olsen, and Christian-Emil Ore. "From XML-tagged Acquisition Catalogues to an Event-based Relational Database," *Proceedings of Computer Applications in Archaeology 2004, Prato, Italy* (in print). http://www.edd.uio.no/artiklar/arkeologi/jordal_caa2004.pdf (accessed May 26, 2009).

INTRASIS. http://www.intrasis.com/engelska/index_eng.htm (accessed June 1, 2009).

Lefebvre, Bastien, Xavier Rodier, and Laure Saligny. "Understanding Urban Fabric with the OH_FET Model Based on Social Use, Space and Time," *Archeologia e Calcolatori* 19 (2008): 195–214.

Ore, Christian-Emil, and Oddrun Rangsæter. *Final Report for the Museum Project 2007.* www.muspro.uio.no/ engelsk-omM.shtml (accessed June 1, 2009).