Dipartimento di Informatica
Università degli Studi di Verona

Modeling Time in Archaeological Data: the Verona Case Study

Alberto Belussi
Sara Migliorini

Abstract

Time and space are two important characteristics of archaeological data. As regards to the first aspect, in literature many time dimensions for archaeology have been defined which extend from the excavation time, to the dating of archaeological objects. Standard ISO 19018 describes temporal characteristics of geographical information in terms of both geometric and topological primitives. The first aim of this report is to analyse the applicability of such Standard for representing archaeological data, referring to the model adopted by the city of Verona (Italy) as case study. However, since archaeological dates are often subjective, estimated and imprecise, one of the main lack in the Standard is the inability to incorporate such vagueness in date representation. Therefore, the second contribution of this report is the extension of the Standard in order to represent fuzzy dates and fuzzy relationships among them. Finally, considering the process through which objects are usually manually dated by archeologists, some existing automatic techniques for time reasoning may be successfully applied in this context in order to guide the dating process. For this purpose, the last report contribution regards the translation of some archaeological temporal data into a Fuzzy Temporal Constraint Network (FTCN) for checking the overall data consistency and reducing the vagueness of some dates based on their relationships with other ones.

Keywords: archaeological data, vague time dimensions, fuzzy temporal knowledge

1 Introduction

Archaeological data are usually managed through Geographical Information Systems (GISs), since one of their main characteristic is an absolute or relative location in 3D space. This information allows to derive important relations between finds, in particular as concern to stratigraphic analysis. Anyway, besides to spatial location, temporal dimension is of considerable interest for archaeological research. For this reason, some attempts can be found in literature in order to define a 4D GIS tailored for archaeological data [5], where the fourth dimension is the temporal one.

Standard ISO 19108 [7] describes the temporal characteristics of geographical information. In particular, it applies the concepts of geometry and topology, which are typical of the spatial domain, to the description of temporal aspects. The main observation is that a point in time occupies a position in a temporal reference system and can be connected with other points through ordering relations. Topological structures can be used to explicitly describe relations between time points, even when they cannot be directly derived. These structures are particularly useful in the archaeological domain, where precedence relations between objects are frequently better known than their location in time. For all these reasons, the first contribution of the report is an evaluation of the Standard applicability for modeling archaeological data. This evaluation is done by considering an existing information system, called SITAVR (Sistema Informativo Territoriale Archeologico di Verona), which collects and manages the archaeological data of Verona, a city in northern Italy [4]. SITAVR development has been started in 2012 through the collaboration with the Archaeological Agency of Veneto Region and a cooperation agreement with the Archaeological Special Agency of Rome, which was developing an information system for the Italian capital since 2007.

From this preliminary evaluation has emerged that the time dimensions of archaeological data are typically vague. Due to this inherent vagueness, many dates are wrongly described as periods instead of instants with the aim to provide a possibility interval for its value. For instance, the construction date of a building can be located between 1830 and 1850 plus-minus 10 years. This kind of date specification suggests the use of a fuzzy approach for representing time. Moreover, also the ordering relations between time points can incorporate a certain degree of possibility. Therefore, the second contribution of the paper is the extension of the Standard with fuzzy types and relations, and the application of them to the SITAVR model.

Time knowledge about particular finds and relations among them are typically used by archaeologists to derive new knowledge or during the in-

terpretation process. In literature many techniques have been proposed for automatically deriving new knowledge from available data. One of this technique is based on the construction of the so called Temporal Constraint Network (TCN) and an extension to the fuzzy context have been developed in [12], called Fuzzy Temporal Constraint Network (FTCN). FTCNs usually incorporate only metric (geometric) information about time, in particular the distance between two time points. Anyway, as stated before, in the archaeological context, logical (topological) information is also of particular interest. Therefore, this paper considers the approach proposed in [2] for integrating quantity and quality temporal information into a FTCN. The third contribution of this report is the translation of information represented using the fuzzy extension of the Standard into a FTCN.

Reasoning techniques on FTCN allow one to answer two main questions: check the network consistency and compute the minimal network in order to reduce some vagueness. The answers to these two questions can be used to guide archeologists in the complex dating and interpretation process. At the end of the report, a portion of the SITAVR information source is translated into a FTCN and some examples of knowledge derivation are provided.

The overall aim of this report is to propose a framework based on existing standards and consolidated reasoning techniques, for representing and managing temporal dimensions of archaeological data. Such framework can become an invaluable help for archeologists during the dating and interpretation processes, and can be applied in other contexts with similar characteristics, such as geology. The remainder of the report is organized as follows: Sec. 2 summarizes some results about the representation of time in archaeology and the fuzzy temporal reasoning. Sec. 3 presents the Standard ISO TC 211 19108, while Sec. 4 applies its concepts to the modeling of the SITAVR information system. Sec. 5 provides detailed background notions about the representation of uncertain time through FTCN considering both quantitative and qualitative information. Sec. 6 describes how the Standard can be extended in order to incorporate fuzzy time dimensions, such extended concepts are used in Sec. 7 for representing SITAVR concepts. Sec. 8 formalizes how the concepts of the extended Standard can be translated into a FTCN. An example of reasoning performed on a real-world case is provided in Sec. 9. Finally, Sec. 10 summarizes the obtained results and discusses future work.

2   Related Work

A first investigation about the applicability of Standard ISO 19108 for the representation of archaeological data is proposed in [5]. The authors conclude

that the standard can be successfully applied in this context, but they also highlight the lack of constructs for describing the inherent vagueness of such data. In [10] the authors identify six potential time categories for archaeological finds which includes: excavation time, database time, stratigraphic time, archaeological time, site phase time and absolute time. The SITAVR model considered in this report includes many of this time categories. In particular, it includes the excavation time, the stratigraphic time (in terms of relative temporal positions between finds), the archaeological time (e.g. Roman Time or Middle Age), the site phase time (i.e. the distinction of different phases during an object life), and the absolute time. In [11] the authors discuss the possibility of incorporating a fuzzy approach into a particular spatio-temporal processing framework in which temporal information is stored through a series of snapshots associated to particular instants in time and relationships regard the relative ordering among events. In this framework spatial objects are temporally located into a specific time layer (snapshot) associated to a particular instant in time. The authors define the concept of fuzzy time layer which is an imprecise time interval within initial and final time points and possibility distribution functions. The proposed model is applied to a wildlife migration modeling analysis.

A Temporal Constraint Network (TCN) [6] is a formalism for representing temporal knowledge based on metric temporal constraints. It supports the representation of temporal relations and is provided with efficient algorithms based on CSP (Constraint Satisfaction Problem) techniques. Recently, a generalization based on fuzzy sets has been proposed in literature, in order to represent vague and unprecise temporal relations. Such extension is known as Fuzzy Temporal Constraint Network (FTCN) [12]. Moreover, in [2] the authors propose a way to integrate quantitative and qualitative relations in a FTCN. In particular, as qualitative relations they consider the well-known Allen's interval algebra [1] and they define a set of functions to transform a qualitative constraint into a quantitative one, and vice-versa. These ideas are further developed in [3] by the same authors, in order to provide a complete fuzzy interval algebra, called IA^{fuzz}. This report considers the work in [2] during the transformation of the temporal knowledge available in a SITAVR model into a FTCN.

3 Standard ISO 19108

Standard ISO 19108 [7] describes the temporal characteristics of geographical information. The schema consists of two packages: `Temporal Objects` and `Temporal Reference System`. Package `Temporal Objects` defines tempo-
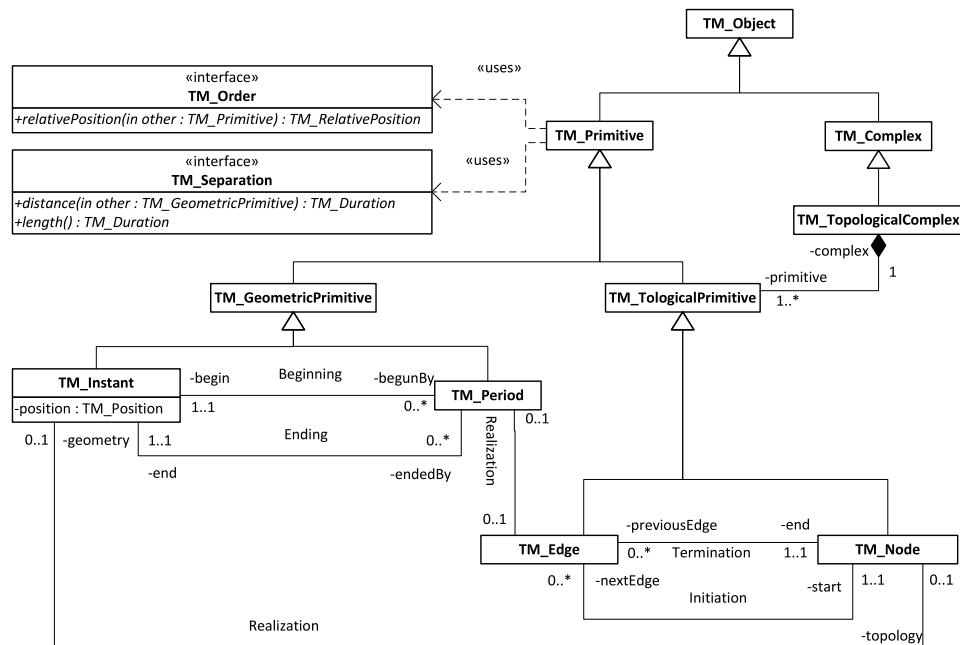
Figure 1: Package `Temporal Object` of the Standard ISO 19108.

ral geometric and topological primitives that shall be used as values for the temporal characteristics of features and datasets. The temporal position of a primitive shall be specified in relation to a temporal reference system. For this purpose, package `Temporal Reference System` provides elements for describing temporal reference systems.

Package `Temporal Objects` is illustrated in Fig. 1. It includes primitive and complex objects: `TM_Primitive` is an abstract class that represents a non-decomposable element of time geometry (`TM_GeometricPrimitive`) or topology (`TM_TopologicalPrimitive`), while `TM_TopologicalComplex` is an aggregation of connected `TM_TopologicalPrimitive`s. Similarly to the corresponding spatial concepts, `TM_GeometricPrimitive` provides information about temporal positions, while `TM_TopologicalPrimitive` provides information about connectivity in time. Both `TM_Primitive`s has a dependency on the interfaces `TM_Order` and `TM_Separation`: the first one provides an operation for determining the relative position of this primitive with respect to another one, while the second interface provides operations for computing the length (duration) of this primitive and the distance from another one.

In the temporal context there are two geometric primitives: *instant* (`TM_Instant`) and *period* (`TM_Period`), and two corresponding topological primitives: *node* (`TM_Node`) which can be realized as an instant, and *edge* (`TM_Edge`) which can be realized as a period.

Each edge starts and ends in nodes, while a node can also exists without being associated with edges. When a node has a realization on the time axis as instant, its temporal position is determined, otherwise it can be qualitative described by means of the temporal relations represented by the edges that starts and ends in the node. When an edge has a realization on the time axis as period, its temporal position is determined, otherwise it simply represents a temporal relation between two nodes and its corresponding period can be only qualitative described by means of its start and end nodes.

In order to deal with a connected set of nodes and edges, the `TM_Topologi-calComplex` class of objects has been introduced. A topological complex is a set of connected topological primitives. Each edge of a topological complex has start and end nodes inside the complex. It can be represented as a graph in which a set of `TM_Primitive`s are contained where the above described constraint on edges is satisfied. A `TM_TopologicalComplex` allows to compactly represent relations among objects. In particular, Allen's relations [1] can be derived as illustrated in Table 1.

Table 1: Allen's temporal relations that can be derived from the structure of a topological complex $\mathcal{C}$. In the table $\mathcal{I}$ is the set of *Initiation* associations and $\mathcal{T}$ is the set of *Termination* association in the topological complex. In the sequel, $(a, b) \in \mathcal{I}$ stands for $a \in$ `TM_Node` $\wedge$ $b \in$ `TM_Edge` $\wedge$ there exists an *Initiation* association between them. A similar definition holds for $(a, b) \in \mathcal{T}$.

Relation	Types	Condition
a before b	any	$\exists$ a sequence $S \in \mathcal{C}$ such that: in S a is earlier than b $\wedge$ $(a,b) \notin \mathcal{I} \wedge (b,a) \notin \mathcal{I} \wedge$ $(a,b) \notin \mathcal{T} \wedge (b,a) \notin \mathcal{T}$.
a meets b	(`TM_Edge`,`TM_Edge`)	$\exists n \in$ `TM_Node`$((a,n) \in \mathcal{T} \wedge (n,b) \in \mathcal{I})$
a begins b	(`TM_Node`,`TM_Edge`)	$(a,b) \in \mathcal{I}$
a begunBy b	(`TM_Edge`, `TM_Node`)	$(b,a) \in \mathcal{I}$
a equals b	any	a and b are the same primitive
a ends b	(`TM_Node`,`TM_Edge`)	$(a,b) \in \mathcal{T}$
a endedBy b	(`TM_Edge`,`TM_Node`)	$(b,a) \in \mathcal{T}$
a metBy b	(`TM_Edge`,`TM_Edge`)	$\exists n \in$ `TM_Node`$((n,a) \in \mathcal{I} \wedge (n,b) \in \mathcal{T})$
a after b	any	$\exists$ a sequence $S \in \mathcal{C}$ such that: in S a is later than b $\wedge$ $(a,b) \notin \mathcal{I} \wedge (b,a) \notin \mathcal{I} \wedge$ $(a,b) \notin \mathcal{T} \wedge (b,a) \notin \mathcal{T}$

Package `Temporal Reference Systems` is illustrated in Fig. 2. Standard ISO 8601 specifies the use of the Gregorian Calendar and 24 hour local, or

Coordinate Universal Time (UTC) for information interchange. This last one shall be the primary temporal reference system for geographical information. Anyway, when more than one temporal reference system is used in a single feature catalogue, application schema, or dataset, the definition of each temporal characteristic shall identify the temporal reference system that is used. Package `Temporal Reference Systems` includes three common types of temporal reference systems: calendars (used with clocks for greater resolution), temporal coordinate systems, and ordinal temporal reference systems.

A `TM_Calendar` is a discrete temporal reference system that provides a basis for defining temporal position with a resolution up to one day. A calendar has an hierarchical structure in which a specific type of time interval is used at each level. In other words, the number and types of temporal granularities provided by a calender depends on the number and types of provided hierarchical levels. A `TM_Clock` can be used with a calendar in order to provide a complete description of a temporal position within a specific day.

A `TM_CoordinateSystem` is a temporal coordinate system based on a continuous interval scale defined in terms of a single time interval: all dates are defined as a multiple of the standard interval associated with the reference system and with respect to a chosen origin. It eases the computation of distances between points and the description of temporal operations which can be complicated when temporal positions are described in terms of calendar dates and times in a day.

A `TM_OrdinalReferenceSystem` is based on an ordinal scale. In its simplest form, it is an ordered series of events. It is particularly appropriate in a number of applications of geographic information (e.g., geology and archeology) in which relative position in time is known more precisely than duration. In such applications, the order of events in time can be well established, but the magnitude of the intervals between them cannot be accurately determined. An ordinal temporal reference system consists of a set of `TM_OrdinalEra`s. They can be often hierarchically structured such that an ordinal era at a given level of the hierarchy includes a sequence of coterminous shorter ordinal era.

A `TM_Position` is a union class that consists of one of the data types listed as its attribute. In particular, it can be a `TM_TemporalPosition` when the application requires to explicitly define the adopted reference system.

A `TM_CalDate` is a data type that shall be used to identify temporal position within a calendar. It has a fundamental property `calDate` which provides a sequence of positive integers where: the first one identifies a specific instance of the unit at the highest level of the calendar hierarchy, the second one identifies a specific instance of the unit used at the next lower level in the hierarchy, and so on. The format defined by Standard ISO 8601 [8]
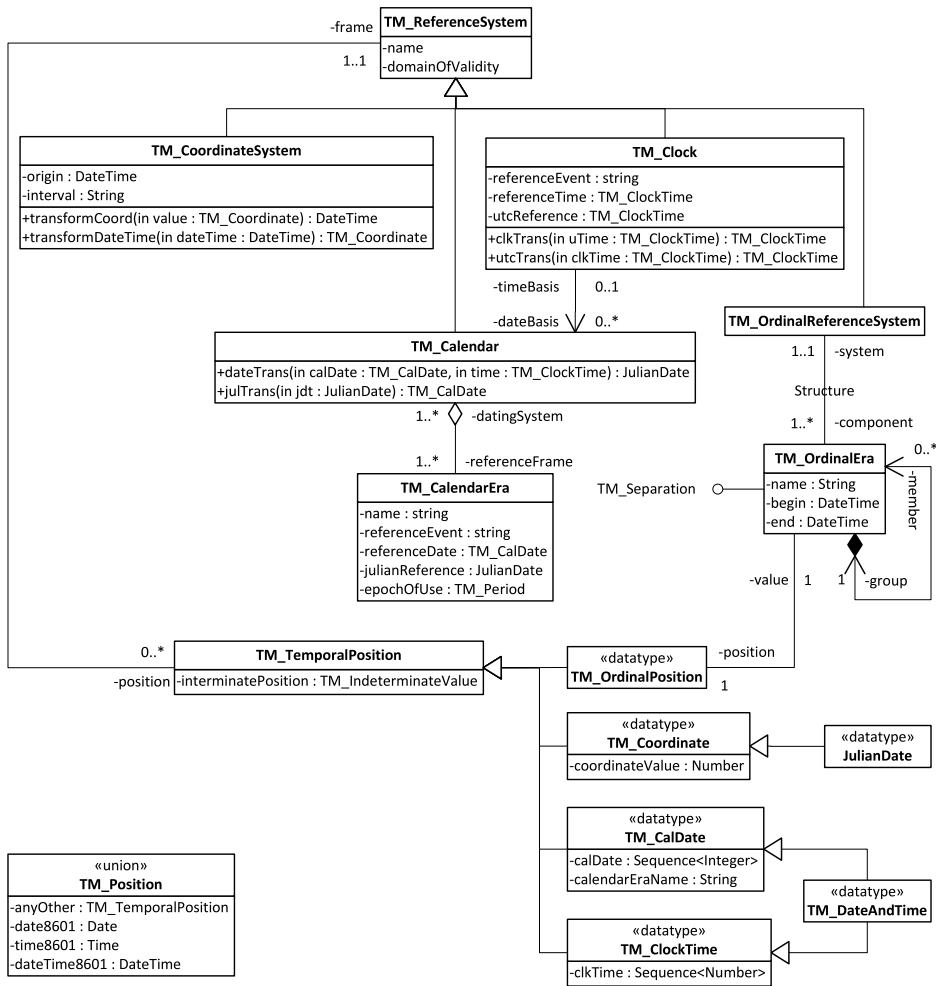
Figure 2: Package `Reference Systems` of the Standard ISO 19108.

for dates in Gregorial calendar may be used for any date that is composed of values for year, month and day. In particular, besides to the complete format `YYYY-MM-DD` which requires to specify a four-digit year, a two-digit month of the year, and a two-digit day of that month, the standard allows to specify dates at reduced precision: `YYYY-MM` which refers to a particular month of a year, `YYYY` which refers to a particular year, and `YY` which refers to a particular century (e.g., 19 stands for the century from 1900 to 1999). To represent years before 0000, the standard also permits an expanded year representation `[± YYYYY]` which uses an extra digits beyond the four-digit minimum: each year must be prefixed with a + or - sign instead of the common AD or BC notation; by convention 1 BC is labelled +0000, 2 BC is labeled -0001, and so on [8].

A `TM_ClockTime` is a data type that shall be used to identify a temporal position within a day. Similarly, a `TM_DateAndTime` is a subclass of both `TM_CalDate` and `TM_ClockTime` which provides a single data type for identifying a temporal position with a resolution of less than a day. A `TM_Coordinate` is a data type that shall be used for identifying temporal position within a temporal coordinate system. A specialization of this type, is `JulianDate` which identifies a position with respect to the Julian proleptic calendar. Finally, `TM_OrdinalPosition` is a data type that shall be used for identifying temporal position within an ordinal temporal reference system.

4 Modeling SITAVR Time Dimensions with Standard ISO 19108

In the archaeological context time dimension may be specified using different reference systems and different calendars. For this reason, this report considers only `TM_TemporalPosition` objects as possible instances for `TM_Position`, while it does not consider `DateTime`, `Date`, or `Time`. In other words, it assumes that the reference system and the used calendar are always explicitly declared.

In SITAVR three main objects of interest can be recognized: `ST_InformationSource`, `ST_ArchaeoPart` and `ST_ArchaeoUnit`, which are also characterized by some temporal dimensions discussed in this section.

An `ST_ArchaeoUnit` is a complex archaeological entity obtained from an interpretation process performed by the responsible officer. Such interpretation is performed based on some finds retrieved during an excavation process or a bibliographical analysis, which are represented by `ST_ArchaeoPart` instances that are selected during the interpretation process. Given an `ST_ArchaeoUnit` object a set of possible temporal *phases* of its evolution are identified, then the selected `ST_ArchaeoPart` objects are assigned to one of the phases. This assignment process is one of the fundamental tasks in archaeology [10]. For instance, examples of phases in the existence of an archaeological entity are: installation/foundation, life/use, and renovation/reuse. In SITAVR the sequence of phases describing the evolution of an `ST_ArchaoeUnit` object is defined as an `ST_Sequence` object, which in turn is a composition of `ST_Phase` objects.

In order to link these classes of the SITAVR model to the ISO Standard [7] we can observe that, since the relative order between each pair of phases is typically known with more certainty than their absolute position, the collection of phases in the sequence of an `ST_ArchaeoUnit` object can
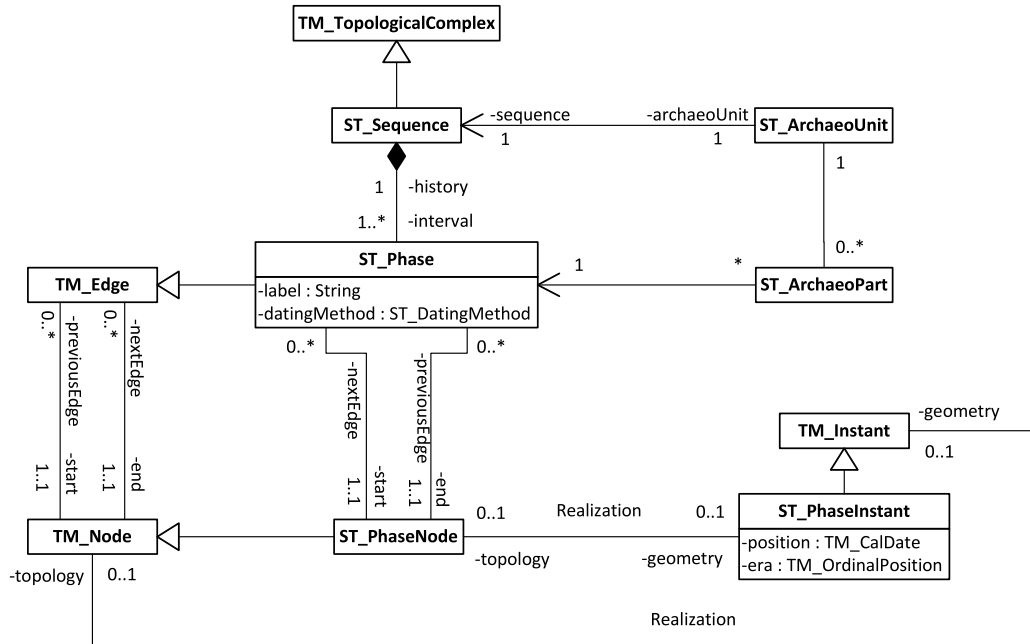
Figure 3: Representation of the time aspects characterizing an archaeological unit in SITAVR.

be modeled using a topological approach, as also suggested in the Standard. Fig. 3 illustrates the final result obtained by applying this solution. More specifically, the `ST_Sequence` object of an `ST_ArchaeoUnit` can be described as a topological complex, thus the `ST_Sequence` class can be declared in the model as a specialization of the `TM_TopologicalComplex` class of the Standard. Moreover, an `ST_Sequence` object is composed of `ST_Phase` objects; therefore, the `ST_Phase` class has to be declared in the model as a specialization of `TM_TopologicalPrimitive` class (i.e., `TM_Edge`, since it represents a period). SITAVR adds two additional properties: a meaningful label (e.g., foundation, use, etc.) and the specification of the dating method (e.g., stratigraphic analysis). Also the `Initiation` and `Termination` associations are specialized, because they connect a `ST_Phase` object with particular nodes (instances of the class `ST_PhaseNode` specializing `TM_Node`) which can be realized with a specialization of `TM_Instant`, called `ST_PhaseInstant`. Each `ST_PhaseInstant` has two attributes: a `position` (inherited from `TM_Instant`), which here can be only of type `TM_CalDate`, and a new attribute, called `era`, which is a `TM_OrdinalPosition`; at least one of them has to be not null. The value of the `era` attribute is a `TM_OrdinalEra` object defined with reference to a particular `TM_OrdinalReferenceSystem`, which is called `ST_NamedYearRange` in SITAVR and is exemplified in Fig. 4.
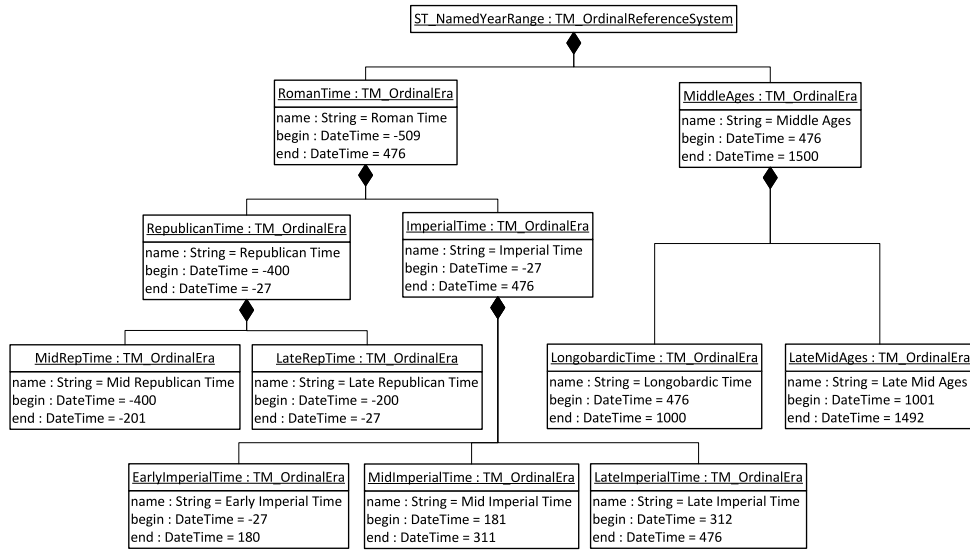
9

Figure 4: Examples of ordinal eras used in SITAVR.

Each ST_ArchaeoUnit is connected to one or more constituent ST_ArchaeoPart, each one representing a single result of an excavation or other investigation processes denoted by the associated information source. For instance, it can be a structural element, a mobile element or a reused element. Each ST_ArchaeoPart is dated in some way and is assigned to a certain phase of the associated ST_ArchaeoUnit. In particular, any ST_ArchaeoPart is characterized by a LifeStartDate role which identifies the beginning of the object life. Moreover, if the partition identifies a structural element, it has also a BuildingDate role which denotes the date of its construction com-
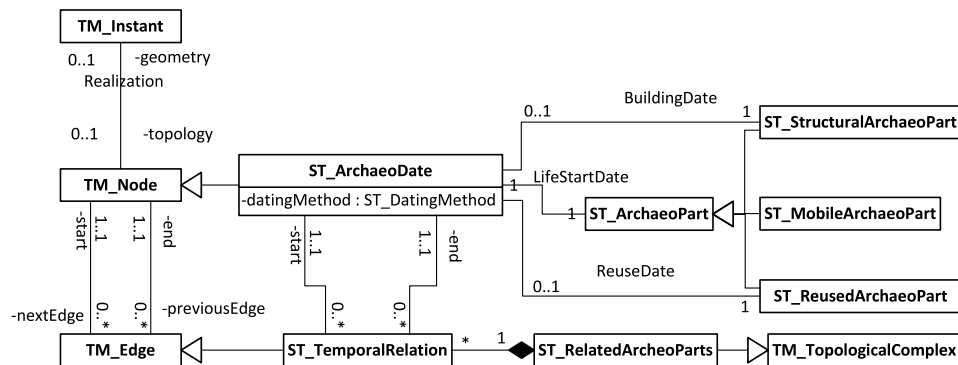
Figure 5: Representation of the time aspects characterizing an archaeological partition in SITAVR.

pletion, while if the partition is a reused element, it is also characterized by a `ReuseDate`. An implicit constraint exists between the life-start date assigned to an archaeological partition and the possible additional dates: both `BuildingDate` and `ReuseDate` have to be after the `LifeStartDate`. Moreoever, constraints can also be defined regarding the phase assignment related to the association of the partition with an archaeological unit: for instance, the life-start date of a mobile or structural partition shall be contained in the assigned phase, while those of a reused element shall precede the phase start node.

The date assigned to an `ST_ArchaeoPart` object is described in the model by the `ST_ArchaeoDate` class, called partition *chronology*, which is related to the ISO Standard as it is a specialization of the `ST_Node` class and has consequently a realization in the `ST_Instant` class. An additional attribute describing the applied dating method characterizes the `ST_ArchaeoDate` class. Exploiting the ISO classes, the chronology of a partition can also be represented by topological primitives, since a relative order between related partitions is better known, than their absolute location. Some edges, called `ST_TopologicalRelation`, can be placed between nodes representing `ST_ArchaeoDate` objects, in order to define temporal relations between related archaeological partition dates. A set of temporal relations related to some connected partitions constitute a topological complex, called `ST_RelatedArchaeoParts`. In accordance with the Standard [7], the relative positions of two `TM_TopologicalPrimitive`s depend upon the positions they occupy within the sequence of `TM_TopologicalPrimitive`s that make up a `TM_TopologicalComplex`, as discussed in Table 1 of Sec. 3.

The following example illustrate a possible topological structure composed of a set of related archaeological partitions.

Example 4.1. *Let us consider four archaeological finds labeled as f_1, f_2, f_3 and f_4 which are coarsely dated as follows: f_1, f_2 are located in the 19th century, while f_3 is dated 1850 and f_4 is dated 1820. Besides these geometrical values, the following temporal relations have been detected: f_1 before f_2 and f_3, while f_2 before f_3 and after f_4. This knowledge can be represented by the topological complex in Fig. 6. Dates associated to nodes f_3 and f_4 are realized as the years 1850 and 1820, respectively. Conversely, dates related to nodes f_1 and f_2 are not realized, but they are located between two dummy nodes representing the years 1800 and 1899. Given such topological structure some automatic reasoning techniques can be applied in order to realize also such dates. In particular, all dates between 1820 and 1850 could be consistent realizations for f_2, while all dates between 1800 and 1820 could be consistent realizations for f_1.* □
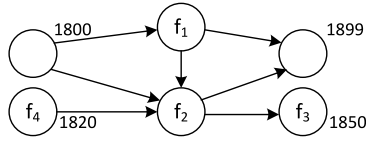
11

Figure 6: Example of topological complex representing ordinal temporal relations between chronologies of archaeological partition.

Each `ST_ArchaeoPart` and each `ST_ArchaeoUnit` refers to an instance of `ST_InformationSource`. An `ST_InformationSource` represents the way used to start collecting information about an archaeological object: for instance, it can be an excavation, a bibliographical study, a construction work, and so on. Each `ST_InformationSource` is characterized by a time dimension that, in accordance to [5], is represented as a geometric primitive, since it is a generally known and documented in some way, as illustrated in Fig. 7. This geometric primitive can be instantiated with both a `TM_Instant` or a `TM_Period` depending on the particular type of information source and the available information.
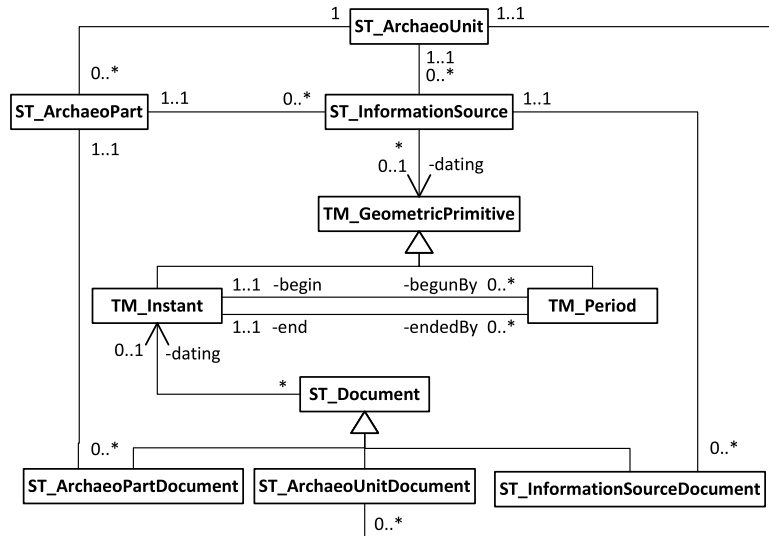
Figure 7: Representation of the time aspects characterizing an information source and a document in SITAVR.

Finally, each of these three main entities can be connected with another SITVAR object characterized by a time dimension: `ST_Document`, which represents a generic collected document; for instance, an excavation report, a cartographic product, and so on. Three specialization of documents are defined: one for information sources, one for archaeological partitions and one

for archaeological units. Similarly to the time dimension of an information source, the dating of this object is also represented by a geometric primitive, since it refers to a modern-age date which is usually well-documented.

5 Representing Vagueness in Time

Several proposals can be found in literature about the representation of temporal knowledge and some reasoning algorithms have been defined for automatically deriving new information. In particular, Temporal Constraint Network (TCN) [6] is a formalism for representing temporal knowledge based on *metric* constraints among pairs of time-points. This report considers only binary constraints, since their expressiveness is satisfactory for many applications. Basic notions about TCN are presented in Sec. 5.1.

However, in the archaeological domain, temporal knowledge is generally characterized by a level of vagueness and dates are usually expressed as interval of great possibility together with a less possible interval. For instance, the construction date of a building can be expressed as: between 1830-1850 plus or minus 10 years. Therefore, a fuzzy representation of time seams to be the more appropriate solution. A generalization of TCN based on fuzzy sets has been proposed in literature [12] in order to cope with vagueness in temporal relations and is presented in Sec. 5.3.

5.1 Temporal Constraint Networks

Definition 5.1 (temporal constraint network)**.** A *temporal constraint network* $\mathcal{N}$ is a tuple $\langle \mathcal{X}, \mathcal{C} \rangle$, where $\mathcal{X}$ is a set of variables representing time points, and $\mathcal{C}$ is a set of binary constraints on those variables. Variables take values on $\mathbb{R}$, while a constraint C_{ij} restrict the duration of the time elapsed between two temporal variables $x_i, x_j \in \mathcal{X}$. □

Definition 5.2 (temporal constraint)**.** Given a TCN $\mathcal{N} = \langle \mathcal{X}, \mathcal{C} \rangle$, a constraint $C_{ij} \in \mathcal{C}$ is represented as:

$$C_{ij} = \{[a_{ij}^1, b_{ij}^1], \ldots, [a_{ij}^k, b_{ij}^k], \ldots, [a_{ij}^n, b_{ij}^n]\}$$

where $a_{ij}^h, b_{ij}^h \in \mathbb{R}$ are values belonging to the variable domain. This constraint states that the temporal distance between variables x_i and x_j is restricted by the following disjunction of inequalities:

$$[a_{ij}^1 \leq x_j - x_i \leq b_{ij}^1] \vee \cdots \vee [a_{ij}^k \leq x_j - x_i \leq b_{ij}^k] \vee \cdots \vee [a_{ij}^n \leq x_j - x_i \leq b_{ij}^n]$$

□

Definition 5.3 (solution). Given a TCN $\mathcal{N} = \langle \mathcal{X}, \mathcal{C} \rangle$ where $\mathcal{X} = \{x_1, \ldots x_n\}$, a tuple $S = (s_1, \ldots, s_n)$ is a *solution* of the constraint network if the assignment $\{x_1 \leftarrow s_1, \ldots, x_n \leftarrow s_n\}$ satisfies all the constraints in $\mathcal{C}$. $\qquad\square$

Definition 5.4 (feasible value). Given a TCN $\mathcal{N} = \langle \mathcal{X}, \mathcal{C} \rangle$ where $\mathcal{X} = \{x_1, \ldots x_n\}$, a value v is a *feasible value* for the variable x_i if there exists a solution S in which $x_i \leftarrow v$. $\qquad\square$

Definition 5.5 (consistent network). A TCN $\mathcal{N} = \langle \mathcal{X}, \mathcal{C} \rangle$ is *consistent* if at least one solution exists. $\qquad\square$

Definition 5.6 (equivalent network). Two networks are equivalent if they have the same variable set and they represent the same solution set. $\qquad\square$

Definition 5.7 (minimal network). A TCN $\mathcal{N} = \langle \mathcal{X}, \mathcal{C} \rangle$ is *minimal* if it is tighter than any other equivalent network. $\qquad\square$

Definition 5.8 (simple network). A constraint network $\mathcal{N} = \langle \mathcal{X}, \mathcal{C} \rangle$ is *simple* if every constraint can be expressed as a single interval. $\qquad\square$

For simple networks the computation of minimal networks and other interesting queries, such as consistency checking, can be computed in polynomial time, while in the general case consistency checking is an NP-hard problem.

A TCN can be represented by means of a directed graph in which each node is associated with a variable and each arc corresponds to the binary constraint between the connected variables.

5.2 Introduction to Fuzzy Set

Fuzzy sets are sets whose elements have degrees of membership.

Definition 5.9 (fuzzy set). A *fuzzy set* F is a pair (U, μ), where U is a set and μ is a *membership function* $\mu : U \to [0, 1]$, such that for all $u \in U$ the value $\mu(u)$ represents the *grade of membership* of u in F. $\qquad\square$

In particular, $\mu(u) = 1$ reflects full membership of u in F, while $\mu(u) = 0$ express the absolute non-membership in F.

Definition 5.10 (core and support). Given a fuzzy set F with membership function $\mu : U \to [0, 1]$, the *core* of F is the crisp set $\{u \in U \mid \mu(u) = 1\}$, while the *support* of F is the crisp set $\{u \in U \mid \mu(u) > 0\}$. $\qquad\square$

This report considers only trapezoidal distributions, since they are sufficiently expressive in practical contexts, while computationally less expensive than general semi-convex functions [2].

A trapezoidal membership function can be encoded by a 4-tuple (a, b, c, d), where the intervals $[b, c]$ and $[a, d]$ represent the core and the support of the fuzzy set, respectively.

5.3 Fuzzy Temporal Constraint Networks

A fuzzy temporal constraint network (FTCN) is a generalization of TCN in which a *degree of possibility* is associated to each possible value of a temporal constraint. In other words, a constraint between a pair of time-points represents a possibility distribution over temporal distances.

Definition 5.11 (fuzzy temporal constraint)**.** Given two temporal variables x_i and x_j, a *fuzzy temporal constraint* C_{ij} between them is represented as a *possibility distribution function* $\pi_{ij} : \mathbb{R} \to [0, 1]$ that constraints the possible values for the temporal distance $x_j - x_i$. $\qquad\square$

In other words $\pi(d)$ represents the possibility degree for the distance $x_j - x_i$ to take the value d under the constraint C_{ij}.

Definition 5.12. (fuzzy temporal constraint network) A *fuzzy temporal constraint network* $\mathcal{F} = \langle \mathcal{X}, \mathcal{C} \rangle$ consists of a set of variables $\mathcal{X} = \{x_1, \dots x_n\}$ and a set of fuzzy temporal constraints $\mathcal{C} = \{C_{ij} \mid i, j < n\}$ between them. $\qquad\square$

As stated in the previous section, this report considers only trapezoidal distributions, which can be represented as a 4-tuple $\langle a, b, c, d \rangle$. In [2] the authors use a richer representation of trapezoidal fuzzy distribution in which the trapeze height can be different from one. More specifically, they introduce a value α_k, called *degree of consistency*, which denotes the height of the trapeze and allows the representation of non-normalized distributions. This report assumes that the initial knowledge produced by archeologists is always represented by a trapeze with height equal to one. However, the conjunction of the given constraints can produce trapezes with an height less than one, as it will be shown in the following; therefore, such parameter cannot be excluded from the constraint formulation. Given such considerations, the notion of fuzzy temporal constraint can be defined as follows.

Definition 5.13 (fuzzy trapezoidal constraint)**.** Given two temporal variables x_i and x_j, a fuzzy trapezoidal temporal constraint $C_{ij} = \{T_1, \dots, T_m\}$ is a disjunction of trapezoidal distributions π_{T_k}, each one denoted by a trapeze $T_k = \langle a_k, b_k, c_k, d_k \rangle [\alpha_k]$, where the characteristics 4-tuple is enriched with a degree of consistency α_k representing its height. $\qquad\square$

The components of a trapeze T_k take values as follows:

- $a_k, b_k \in \mathbb{R} \cup \{-\infty\}$

- $c_k, d_k \in \mathbb{R} \cup \{+\infty\}$

- $\alpha_k \in [0, 1]$

- $supp(\pi_{T_k}) = \{x : \pi_{T_k}(x) > 0\} = [a_k, d_k]$

- $core(\pi_{T_k}) = \{x : \pi_{T_k}(x) = 1\} = [b_k, c_k]$

Definition 5.14 (well-formed trapeze). A trapeze $T = \langle a, b, c, d$ is *well-formed*, if $a \leq b \leq c \leq d$. □

This definition allows several shapes for trapezes, as illustrated in Fig. 8. In particular, when $b_k = c_k$ the core is formed by only one value and the distribution is also known as triangular function. The set of well-formed trapeze T_k will be denoted as $\mathcal{T}$.
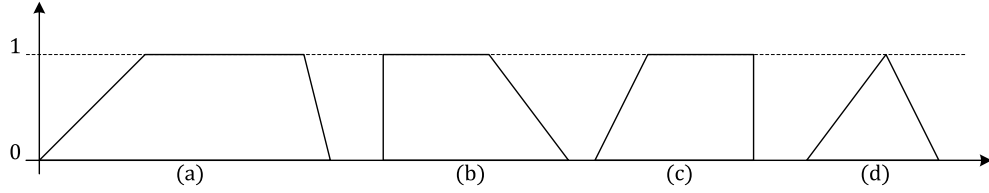
Figure 8: Possible shapes of a trapezoidal possibility distribution function: (a) $a < b < c < d$, (b) $a = b < c < d$, (c) $a < b < c = d$, and (d) $a < b = c < d$.

The semantics of a constraint $C_{ij} = \{T_1, \ldots, T_m\}$ is the possibility distribution function $\pi_{C_{ij}}$ corresponding to the disjunction of the trapezoidal distribution $\pi_{T_k} : \mathbb{R} \to [0, 1]$ for $k = 1, \ldots, m$.

Definition 5.15 (trapezoid possibility distribution function). The *possibility distribution function* of a generic trapeze $T_k \in \mathcal{T}$ can be written as:

$$
\pi_{T_k}(x) = \begin{cases} 0 & \text{if } x < a_k \lor x > d_k \\ \alpha_k \cdot \left(\dfrac{x - a_k}{b_k - a_k} \right) & \text{if } a_k \leq x < b_k \\ \alpha_k \cdot \left(\dfrac{x - d_k}{c_k - a_k} \right) & \text{if } c_k < x \leq d_k \\ \alpha_k & \text{otherwise} \end{cases}
$$

□

Definition 5.16 (solution). Let $\mathcal{F} = \langle \mathcal{X}, \mathcal{C} \rangle$ be a fuzzy temporal constraint network. An n-tuple $S = \{s_1, \ldots s_n\}$, where $s_i \in \mathbb{R}$, is a *possible solution* of $\mathcal{F}$ at degree α if and only if:

$$\deg(S) = \min_{i,j}\{\pi_{C_{ij}}(s_j - s_i)\} = \alpha$$

where π_{ij} stands for the possibility distribution associated to the constraint C_{ij} and the degree corresponds to the least satisfied constraint. $\square$

In the case of a FTCN, each solution is characterized by a *degree of satisfaction* reflecting a trade-off among potentially conflicting constraints.

Definition 5.17. (equivalent networks) Two FTCNs are equivalent if they have the same fuzzy set of solutions. $\square$

Definition 5.18. (optimal solution) Given a FTCN $\mathcal{F}$, a solution for $\mathcal{F}$ is optimal if it maximizes its degree of satisfaction. $\square$

The most widely used algorithm for constraint propagation is the *path-consistency algorithm*.

Definition 5.19 (path-consistency algorithm). Given three variables x_i, x_k and x_j of a FTCN $\mathcal{F}$, a new constraint between x_i and x_j can be induced from pre-existing constraints by the path consistency algorithm as follows.

Given a generic local instantiation $x_i = d_i$, $x_j = d_j$, the degree of satisfaction of any solution extending it is limited by $\pi_{C_{ik}}$ and $\pi_{C_{kj}}$, and in particular it cannot be greater than $\pi'_{ij}(d_j - d_i)$, where π'_{ij} is the composition (addition between fuzzy sets) between $\pi_{ik}(x)$ and π_{kj}:

$$\forall x \in \mathbb{R} \ (\pi'_{ij}(x) = \pi_{ik} \circ \pi_{kj}(x) = \sup_{t_1, t_2 \in \mathbb{R}, t_1 + t_2 = x} \{\min\{\pi_{ik}(t_1), \pi_{kj}(t_2)\}\})$$

Since also the pre-existing constraints C_{ij} must be taken into account, then π'_{ij} must be conjuncted with pre-existing relations π_{ij} in order to identify the induced constraint:

$$\forall x \in \mathbb{R} \ (\pi_{ij} \otimes \pi'_{ij}(x) = \min\{\pi_{ij}(x), \pi'_{ij}(x)\})$$

$\square$

The algorithm in the previous definition is exemplified in Fig. 9 where the existing constraint between variables x_i and x_j, represented by the function π_{ij}, is combined with the constraint obtained by adding the existing constraints between $x_i - x_k$ and $x_k - x_j$, represented by the functions π_{ik} and π_{kj}, respectively.
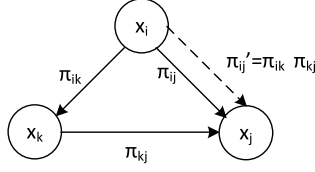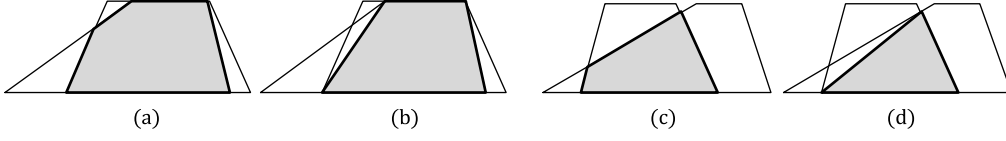
$\pi_{ij}' = \pi_{ik} \circ$

Figure 10: Two examples of approximated conjunction operation $\otimes_a$ between trapezoids: in (a) and (c) the result of the classical conjunction operation between fuzzy possibility distribution functions, and in (b) and (d) the corresponding approximation which produces a trapeze.

is more complex to specialize than composition, because given two generic trapezoids T_1 and T_2, the function $T_1 \otimes T_2 = \min\{T_1, T_2\}$ is not always a trapeze: Fig. 10.a and Fig. 10.c contain two examples of such situation. Therefore, some sort of approximation of $T_1 \otimes T_2$ has to be defined which is a trapeze. For the application context proposed by this report, the following approximation criteria formulated in [2] are appropriate, where T is the result of the approximated conjunction:

- $core(\pi_T) = core(\pi_{T_1} \otimes \pi_{T_2})$

- $h(\pi_T) = h(\pi_{T_1} \otimes \pi_{T_2})$

- $supp(\pi_T) \subseteq supp(\pi_{T_1} \otimes \pi_{T_2})$

In other words, the approximation shall ensure that the core of the obtained distribution is maintained while the possibility of the support elements outside the core can be sightly modified. This operation is formalized as follows.

Definition 5.22 (conjunction $\otimes_a$)**.** Given two constraints C_1 and C_2, the conjunction between two trapezoids $T_1 = \langle a_1, b_1, c_1, d_1 \rangle [\alpha_1] \in C_1$ and $T_2 = \langle a_2, b_2, c_2, d_2 \rangle [\alpha_2] \in C_2$ is defined as follows:

$$T_1 \otimes_a T_2 = T \in \mathcal{T}^{inf}(T_1, T_2) : \forall T_1 \in \mathcal{T}^{inf}(T_1, T_2), \pi_{T_i} \leq \pi_T$$
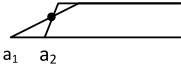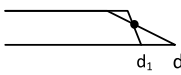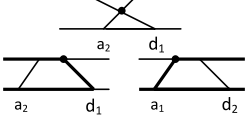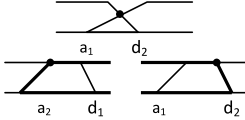
where $\mathcal{T}^{inf}(T_1, T_2) = \{T \mid \pi_T \leq \pi_{T_1} \otimes \pi_{T_2} \ \wedge \ h(\pi_T) = h(\pi_{T_1} \otimes \pi_{T_2})\}$. The trapezoid T can be computed as follows:

$$T = (\max\{a_1, a_2\}, b', c', \min\{d_1, d_2\})[\min\{\alpha_1, \alpha_2\}]$$

where b' and c' depends on the 8 possible intersections between T_1 and T_2 illustrated in Table 2 [2]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\Box$

The set $\mathcal{T}^{inf}$ is the set of trapeze that approximate the conjunction from "below", the result of the conjunction is the greatest trapeze in this set.

Table 2: Possible intersection between two trapezes and corresponding element of the conjunction result.

Situation		Result
a_1 a_2	$a_2 \in (a_1, b_1)$	$b' = \begin{cases} b_1 & \text{if } \alpha_1 = \alpha_2 \wedge b_1 > b_2 \\ b_1 & \text{if } \alpha_1 < \alpha_2 \\ b_2 & \text{otherwise} \end{cases}$
d_1 d_2	$d_1 \in (c_2, d_2)$	$c' = \begin{cases} c_1 & \text{if } \alpha_1 = \alpha_2 \wedge c_1 > c_2 \\ c_1 & \text{if } \alpha_1 < \alpha_2 \\ c_2 & \text{otherwise} \end{cases}$
a_2 d_1 / a_2 d_1 a_1 d_2		b' is the highlighted intersection point.
a_1 d_2 / a_2 d_1 a_1 d_2		c' is the highlighted intersection point.

Some examples of $T_1 \otimes_a T_2$ are illustrated in Fig. 10. In case (d) it is evident that the height of the resulting trapeze can become less than one, hence the degree of consistency α becomes necessary.

The disjunction operation is not required by the path consistency algorithm, but it can be useful for eliminating redundant trapezes that are accidentally introduced by users or are due to constraint propagation. The disjunction of two general fuzzy distribution functions π_1 and π_2 is defined as $\forall d \in \mathbb{R} : \pi_1 \oplus \pi_2(d) = \max\{\pi_1(d), \pi_2(d)\}$. However, like conjunction, disjunction is not closed in the algebra of trapezoids. Therefore, the idea is to compute a tentative trapeze, and then check whether it corresponds to the disjunction of the involved constraints (i.e., correspond of one of the two involved trapezes), otherwise the constraints will be maintained separated.
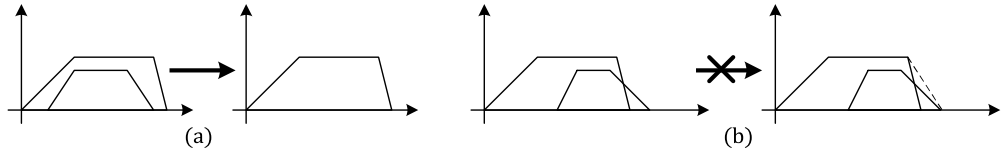

Figure 11: Two examples of approximated disjunction operation $\oplus_a$ between trapezoids: in (a) the operation can be performed, while in (b) the operation cannot be performed.

Definition 5.23 (disjunction $\oplus_a$). Given two constraints C_1 and C_2, the disjunction between two trapezes $T_1 = \langle a_1, b_1, c_1, d_1 \rangle [\alpha_1] \in C_1$ and $T_2 =$

20

$\langle a_2, b_2, c_2, d_2 \rangle[\alpha_2] \in C_2'$ is defined as follows [2]:

$$T_1 \otimes T_2 = \langle a, b, c, d \rangle[\max\{\alpha_1, \alpha_2\}]$$

where:

- $a = \min\{a_1, a_2\}$

- $b = \begin{cases} b_1 & \text{if } \alpha_1 > \alpha_2 \\ b_2 & \text{if } \alpha_2 > \alpha_1 \\ \min\{b_1, b_2\} & \text{otherwise} \end{cases}$

- $c = \begin{cases} c_1 & \text{if } \alpha_1 > \alpha_2 \\ c_2 & \text{if } \alpha_2 > \alpha_1 \\ \max\{c_1, c_2\} & \text{otherwise} \end{cases}$

- $d = \max\{d_1, d_2\}$ $\qquad \square$

Fig. 11.a illustrates a case where the disjunction is executed, while Fig. 11.b illustrates a case where it cannot be executed.

5.4   Fuzzy Qualitative Temporal Constraints

Qualitative temporal constraints can be represented using the Allen's Interval Algebra [1]. An extension of this model that integrates the ideas of flexibility and vagueness has been presented in [2, 3] and is called IA^{fuz} algebra.

Definition 5.24 (qualitative constraint). A *qualitative constraint* is a binary relation between a pair of intervals I_i and I_j, represented as a disjunction of atomic relations:

$$I_i(rel_1, \dots rel_m)I_j$$

where each rel_k is one of the 13 mutually exclusive atomic relations: *before, after, begins, ends, during, equals, contains, overlaps, meets, overlappedBy, metBy, begunBy, endedBy.* $\qquad \square$

In order to integrate the concept of vagueness and flexibility into Allen's framework, each atomic relation rel_k composing a qualitative constraint is enriched with a degree α_k representing its *preference degree*.

Definition 5.25 (fuzzy qualitative constraint). Given two temporal intervals I_i and I_j, a fuzzy qualitative constraint C_{ij} between them is represented as:

$$C_{ij} = (rel_1[\alpha_1], \dots, rel_{13}[\alpha_{13}])$$

where $\alpha_k \in [0, 1]$ is the preference (membership) degree of rel_k [2]. $\qquad \square$

Operations on fuzzy qualitative constraints can be defined as in the following definitions [2].

Definition 5.26 (inversion). Given a fuzzy qualitative constraint $C = (rel_1[\alpha_1], \ldots, rel_{13}[\alpha_{13}])$, its inversion C^{-1} is defined as:

$$C^{-1} = (rel_1^{-1}[\alpha_1], \ldots, rel_{13}^{-1}[\alpha_{13}])$$

where rel_k^{-1} is the classical operation defined accordingly with the Allen's inversion table [2]. $\square$

Definition 5.27 (conjunction $\otimes$). Given two fuzzy qualitative constraints $C^1 = (rel_1[\alpha_1^1], \ldots, rel_{13}[\alpha_{13}^1])$ and $C^2 = (rel_1[\alpha_1^2], \ldots, rel_{13}[\alpha_{13}^2])$, their conjunction $C = C_1 \otimes C_2$ is defined as $C = (rel_1[\alpha_1], \ldots, rel_{13}[\alpha_{13}])$ where [2]:

$$\alpha_k = \min\{a_k^1, a_k^2\} \ \ k \in \{1 \ldots 13\}$$

$\square$

Definition 5.28 (disjunction $\oplus$). Given two fuzzy qualitative constraints $C^1 = (rel_1[\alpha_1^1], \ldots, rel_{13}[\alpha_{13}^1])$ and $C^2 = (rel_1[\alpha_1^2], \ldots, rel_{13}[\alpha_{13}^2])$, their conjunction $C = C_1 \oplus C_2$ is defined as $C = (rel_1[\alpha_1], \ldots, rel_{13}[\alpha_{13}])$ where [2]:

$$\alpha_k = \max\{a_k^1, a_k^2\} \ \ k \in \{1 \ldots 13\}$$

$\square$

Definition 5.29 (composition $\circ$). Given two fuzzy qualitative constraints $C^1 = (rel_1[\alpha_1^1], \ldots, rel_{13}[\alpha_{13}^1])$ and $C^2 = (rel_1[\alpha_1^2], \ldots, rel_{13}[\alpha_{13}^2])$, their composition $C = C_1 \circ C_2$ is defined as $C = (rel_1[\alpha_1], \ldots, rel_{13}[\alpha_{13}])$ where:

$$\alpha_k = \max_{u,v:rel_k \in \{rel_u \circ rel_v\}} \min\{a_k^1, a_k^2\} \ \ k, u, v \in \{1 \ldots 13\}$$

where $rel_u \circ rel_v$ is the classical operation defined accordingly with the Allen's composition table [2]. $\square$

5.5 Integrating Quantitative and Qualitative Fuzzy Temporal Constraints

This section discusses how to integrate quantitative and qualitative fuzzy temporal constraints presented in Sec. 5.3 and Sec. 5.4, respectively. Accordingly with the Standard ISO 19108, in the model of Sec. 4 qualitative

temporal constraints are represented by topological structures in which temporal primitives are connected through edges. More specifically, each edge denotes a precedence relation between time points.

The 13 Allen's temporal relations considered in the previous section have been originally defined in terms of interval variables, instead of instant variables. Anyway, the Standard identifies which of those relations can be applied also in presence of instant variables, and specifies the possible relations on the basis of the type of the involved variables. The relations that can be expressed by topology have been reported in Table 1.

In [2] the authors defines a new algebra PA^{fuzz} in order to express qualitative knowledge concerning points.

Definition 5.30 (fuzzy qualitative constraint between points)**.** Given two time-points p_i and p_j a fuzzy qualitative constraint C_{ij} between them is defined as follows [2]:

$$C_{ij} = (before[\alpha_1], \, equal[\alpha_2], \, after[\alpha_3])$$

where *before*, *equal* and *after* are the possible qualitative relations between time points, and $\alpha_k \in [0,1]$. $\square$

Given such definition, the authors also define an algebra PI^{fuz} for representing qualitative relations between points and intervals, and an algebra IP^{fuz} for representing qualitative relations between intervals and points. The semantics of all these algebras and the relevant operations can be defined in a similar way.

In order to combine qualitative and quantity fuzzy temporal constrains, it is necessary to define some transformation functions between them. In particular, for the purpose of this report the interesting transformation is the qualitative-to-quantitative one.

Definition 5.31 (qualitative to quantitative)**.** Given a qualitative constraint $C = (before[\alpha_1], \, equal[\alpha_2], \, after[\alpha_3])$ between two time points, its corresponding quantitative constraint can be computed using by the function $quan^{fuz}$ defined as follows:

$$
\begin{cases}
\text{if } \alpha_1 > 0 & \text{then } \langle 0, 0, +\infty, +\infty \rangle [\alpha_1] \in quan^{fuz}(C) \\
\text{if } \alpha_2 > 0 & \text{then } \langle 0, 0, 0, 0 \rangle [\alpha_2] \in quan^{fuz}(C) \\
\text{if } \alpha_3 > 0 & \text{then } \langle -\infty, -\infty, 0, 0 \rangle [\alpha_3] \in quan^{fuz}(C)
\end{cases}
$$

$\square$

6 Extending the Standard for Modeling Vague Time Dimensions

The main lack of the Standard ISO 19108 in the representation of archaeological time is the absence of constructs for expressing vagueness. This section analyses how fuzzy concepts can be incorporated into the model presented in the Sec. 3. In particular, this report concentrate on trapezoidal fuzzy distributions, since they are more simple to represent and manage, as stated in Sec. 5. Moreover, they provide a sufficient representation of the time knowledge generally provided by archeologists. Indeed, they usually specify dates like from 1820 to 1850 plus-minus 10 years.

As a general idea, each possible `TM_TemporalPosition` will be extended in order to express a possibility membership function instead of a certain date. In particular, in the archaeological context time granularity is never more fine-grained than a day, thus we can safely omit to consider the `TM_ClockTime` and `TM_DateAndTime` datatypes. The fuzzy extension of the temporal position is illustrated in Fig. 12.
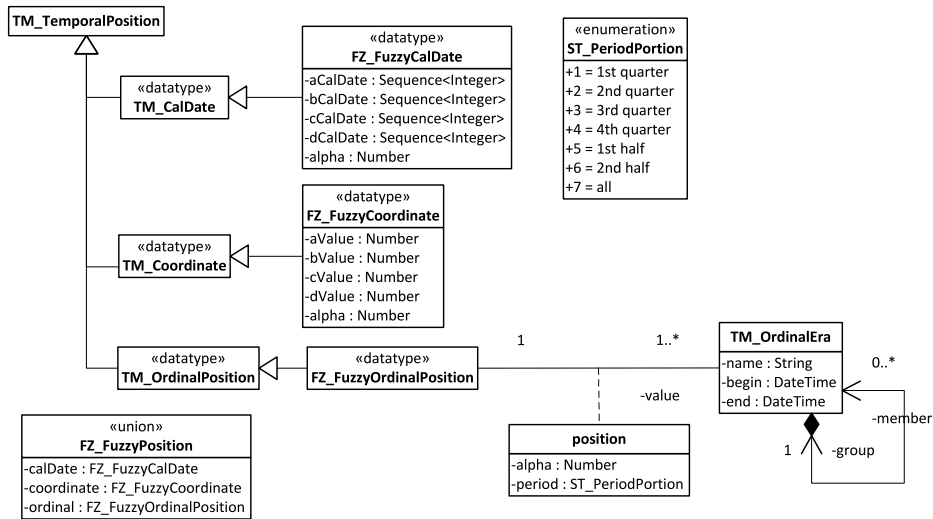

Figure 12: Fuzzy extension of the temporal positions in Standard ISO 19108.

Each calendar date is represented in a fuzzy form using the `FZ_FuzzyCalDate` datatype which contains a trapezoidal tuple $\langle a, b, c, d \rangle [\alpha]$, similar to the one in Sec. 5.3, where a, b, c, d are sequences of integers representing dates.

Similarly, a fuzzy ordinal positions inside an ordinal temporal reference system is represented with a specialized class `TM_FuzzyOrdinalPosition`, which has a qualified association with the corresponding `TM_OrdinalEra` en-

24

riched with a degree of possibility $\alpha \in [0,1]$ and a period. The `period` attribute allows to specify a portion (e.g., the first quarter) of the selected era which is more possible. Moreover, the cardinality on the era side is changed from 1 to 1..∗, since different positions can be defined each one with a possibility value. These positions can be interpreted as a disjunction of positions.

Finally, each coordinate inside a coordinate reference system is extended by the datatype `FZ_FuzzyCoordinate` which contains four numeric values representing the trapeze extremes, and the value α.

These datatypes can be used as value for the union `FZ_FuzzyPosition`. A `FZ_FuzzyPosition` is the type of the `position` attribute of a generic `FZ_FuzzyInstant` which is a fuzzy specialization of a temporal instant.
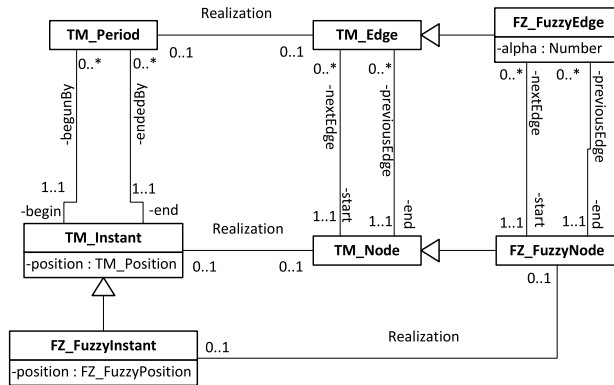

Figure 13: Fuzzy extension of the topological primitives in Standard ISO 19108.

The last aspect to be consider regards the relative ordering between topological primitives inside the same topological complex. In particular, the Standard establishes how to determine the relative ordering between topological primitives, based on their position in the sequence that makes up the topological complex. However, in a fuzzy environment such relations cannot be certain but are characterized by a possibility value. Therefore, a specialization of `TM_Edge` is defined which is called `TM_FuzzyEdge` and is enriched with a possibility value $\alpha \in [0,1]$, as illustrated in Fig. 13. When the a `FZ_FuzzyEdge` is not realized, it simply represents an uncertain relation between two nodes, while when it is realized the corresponding period is characterized by two fuzzy extremes, as illustrated in Fig. 13.

25

7 Modeling Vague SITAVR Time

This section illustrates how the fuzzy datatypes presented in the previous section can be used for modeling vague time aspects in SITAVR.
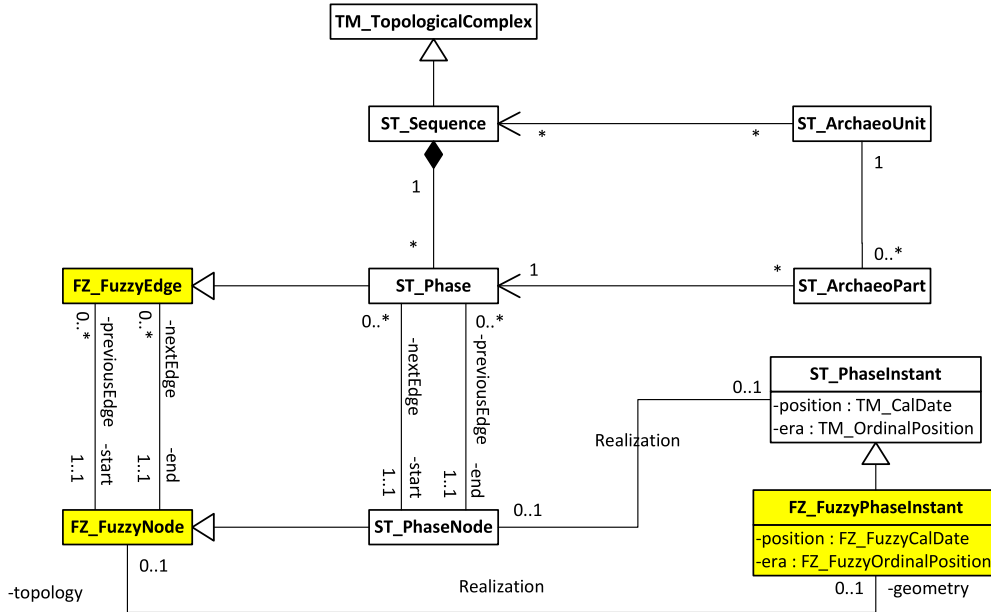

Figure 14: Fuzzy representation of a `ST_ArchaeoUnit`.

As regards to `ST_ArchaeoUnit`, the extension of its time aspects is illustrated in Fig. 14. First of all, a phase instant is represented by a `ST_FuzzyPhaseInstant` whose `position` and `era` attributes are redefined to be typed with the corresponding fuzzy datatypes. Similarly, the topology is represented by the corresponding fuzzy types making uncertain the relation between phases.

An `ST_ArchaeoPart` has three time dimensions: its own dating, the relation with a phase of the corresponding archaeological unit, the definition of a set of time relationships between archaeological partitions. Each `ST_ArchaeoDate` can be realized through a `FZ_FuzzyInstant` in order to express the vagueness of the dating process. Conversely, the assignment to a particular phase remains unchanged even in presence of vagueness, while the relation between two archaeological partitions is represented by a `FZ_FuzzyEdge` in order to assign a possibility value to each relation. This new representation of `ST_ArchaeoPart` is illustrated in Fig. 15.
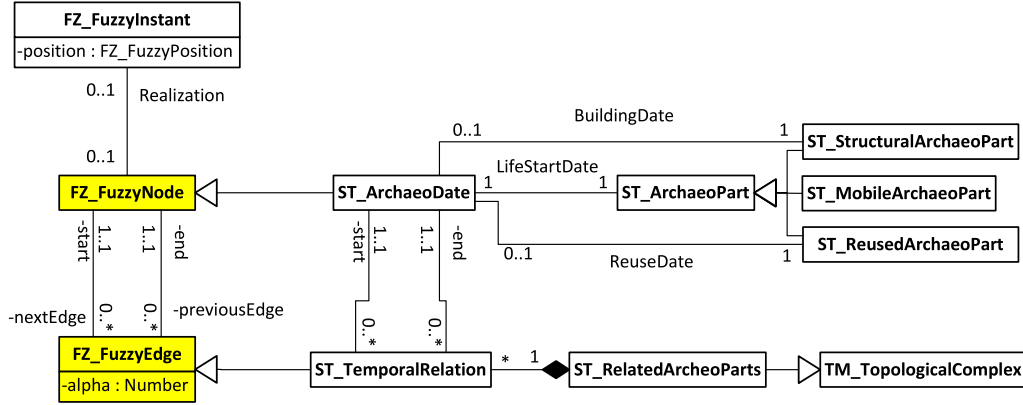
26

Figure 15: Fuzzy representation of a `ST_ArchaeoPart`.

8   Translation of the Fuzzy Model to FTCN

In order to translate the temporal elements introduced in Sec. 6 into a FTCN, it is necessary to initially define a `TM_CoordinateSystem` for transforming all dates into a real number and facilitating the required comparison and operations. The origin of such coordinate system will become the start node of the FTCN and all dates in the network will be defined as multiple of the chosen interval which is the minimum common granularity in the model.

Notice that in a SITAVR model dates can be defined with different granularities: for instance, the components of a fuzzy calendar date can be defined in terms of months or years, not only of days (e.g. the tuple $\langle 1810, 1820, 1850, 1860 \rangle[1]$ is a valid fuzzy date). Nevertheless, all the components of a given date (fuzzy tuple) have the same granularity. The following rule allows to transform all dates in the model to a common minimum granularity.

Rule 8.1 (minimum granularity). Let g the minimum common granularity in the considered model (i.e., day, month or year). Any fuzzy date $x = \langle a, b, c, d \rangle[\alpha]$ whose components have a granularity smaller than g, will be transformed into a date with granularity g in the following way:

- If g is "day" and the granularity of x is "month": the left extremes of the trapeze (i.e. components a and b) become the first day of the given month, while the right extremes of the trapeze (i.e. components c and d) become the last day of the given month.

- If g is "day" and the granularity of x is "year": the left extremes of the trapeze (i.e. components a and b) become the first day of the first

27

month of the given year, while the right extremes of the trapeze (i.e. components c and d) become the last day of the last month of the given year.

- If g is "month" and the granularity of x is "year": the left extremes of the trapeze (i.e. components a and b) become the first month of the given year, while the right extremes of the trapeze (i.e. components c and d) become the last month of the given year.

- All the other combinations does not require any transformation. □

This transformation allows to obtain a trapeze that entirely covers the specified month/year. Clearly, this granularity is useful only for reasoning purposes and does not affect the granularity of the represented knowledge. Given such rule, the following transformations assumes that all dates have been reported to a uniform granularity.

Rule 8.2 (calendar date). A `FZ_FuzzyCalDate` $x = \langle$`aCalDate, bCalDate, cCalDate, dCalDate`$\rangle[\alpha]$ is firstly translated into the tuple $x = \langle a, b, c, d \rangle[\alpha]$ where $a, b, c, d \in \mathbb{R}$ is the representation of `aCalDate, bCalDate, cCalDate, dCalDate` into the chosen coordinate reference system, respectively. Secondly, the tuple x is transformed into the portion of FTCN illustrated in Fig. 16.a, where s is the start node of the network. □
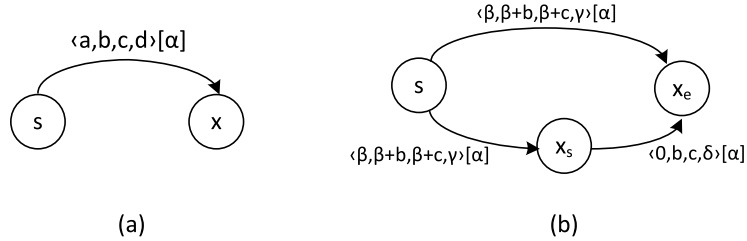
Figure 16: (a) Translation of a fuzzy calendar date to FTCN. (b) Translation of a fuzzy position inside an ordinal era.

Rule 8.3. (coordinate) The translation of a `TM_FuzzyCoordinate` x is similar to the translation of a `TM_FuzzyCalDate`; howerver, it requires an initial transformation of its position to the chosen coordinate reference system only if it is different from the one associated to x. □

Rule 8.4 (ordinal position). Each `TM_FuzzyOrdinalPosition` related to a `TM_OrdinalEra` x is translated into two nodes x_s and x_e, representing the

28

extremes of the era or of its considered portion. These nodes are connected by an arc labeled as in Table 3. Moreover, an arc is added from the start node s to x_s and from s to x_e with a label defined in Table 3. This translation is illustrated in Fig. 16.b.

Table 3: Translation of the relation between a ordinal position and its corresponding era x, where $\beta = x.\texttt{begin}$ and $\gamma = x.\texttt{end}$ are the era boundaries expressed with respect to the considered coordinate reference system, and $\delta = x_e - x_b$ is the era duration.

Period Portion	Arcs $s \to x_s$, $s \to x_e$	Arc $x_s \to x_e$
1st quarter	$\langle \beta, \beta, \beta + \delta/4, \gamma \rangle[1]$	$\langle 0, 0, \delta/4, \delta \rangle[1]$
2nd quarter	$\langle \beta, \beta + \delta/4, \beta + \delta/2, \gamma \rangle[1]$	$\langle 0, \delta/4, \delta/2, \delta \rangle[1]$
3rd quarter	$\langle \beta, \beta + \delta/2, \beta + 3\delta/4, \gamma \rangle[1]$	$\langle 0, \delta, 3\delta/4, \delta \rangle[1]$
4th quarter	$\langle \beta, \beta + 3\delta/4, \gamma, \gamma \rangle[1]$	$\langle 0, 3\delta/4, \delta, \delta \rangle[1]$
1st middle	$\langle \beta, \beta, \beta + \delta/2, \gamma \rangle[1]$	$\langle 0, 0, \delta/2, \delta \rangle[1]$
2nd middle	$\langle \beta, \beta + \delta/2, \gamma, \gamma \rangle[1]$	$\langle 0, \delta/2, \delta, \delta \rangle[1]$
all	$\langle \beta, \beta, \gamma, \gamma \rangle[1]$	$\langle 0, 0, \delta, \delta \rangle[1]$

Rule 8.5 (node). Each not-realized `FZ_FuzzyNode` x is also translated into a node x and connected with an arc $\langle 0, 0, +\infty, +\infty \rangle[1]$ starting from the start network node. $\qquad\square$

Rule 8.6 (edge). Each `FZ_FuzzyEdge` from a `FZ_FuzzyNode` x to a `FZ_Fuzzy-Node` y is translated into an edge from x to y labeled with the constraint $\langle 0, 0, +\infty, +\infty \rangle[1]$. $\qquad\square$

Application Specific Constraints

In the SITAVR model presented in the Sec. 7 different implicit constraints are defined between the time dimensions of object instances. Assuming the existence a logic language like OCL [9] for representing such temporal constraints in the model, this section provides some examples of application specific constraints and their translation to FTCN.

Each `ST_FuzzyPhaseInstant` has two attributes that can be specified: a fuzzy calendar date or a fuzzy position inside an era. In case both attributes are specified, an implicit constraint is considered during the FTCN

construction:

$\forall o \in$ ST_FuzzyPhaseInstant(

$\neg(o.\mathtt{position}.isUndefined() \vee o.\mathtt{era}.isUndefined()) \implies$

$(o.\mathtt{era}.\mathtt{begin}.\mathrm{before}(o.\mathtt{position}) \vee o.\mathtt{era}.\mathtt{begin}.\mathrm{equal}(o.\mathtt{position}))\wedge$

$(o.\mathtt{era}.\mathtt{end}.\mathrm{after}(o.\mathtt{position}) \vee o.\mathtt{era}.\mathtt{end}.\mathrm{equal}(o.\mathtt{position}))))$

As regards to the translation, if the era position is represented by two nodes x_s and x_e connected by an arc $\langle a, b, c, d\rangle[1]$ and the calendar date is represented by a node y, then an arc is added from x_s to y and one from y to x_e with label $\langle 0, 0, c - b, d - a\rangle[1]$.

Moreover, an implicit constraint exists between the dating of each archaeological partition and the associated phase. In particular, let us assume that the phase is translated into an edge between two FTCN nodes n and m connected by an edge $n \xrightarrow{\langle a,b,c,d\rangle[1]} m$. The implicit derived relations are determined by the type of archaeological partition as follows:

- ST_MobileArchaeoPart.
  Constraint: The associated LifeStartDate x shall be contained into the assigned phase.

  $\forall p \in$ ST_MobileArchaeoPart

  $(x = p.\mathtt{LifeStartDate} \wedge$

  $(n.\mathrm{equal}(x) \vee n.\mathrm{before}(x)) \wedge (m.\mathrm{equal}(x) \vee m.\mathrm{after}(x)))$

  which is represented by the following two additional edges:

  $- \ n \xrightarrow{\langle 0,0,d-a,d-a\rangle[1]} x$
  $- \ x \xrightarrow{\langle 0,0,d-a,d-a\rangle[1]} m$

- ST_StructuralArchaeoPart
  Constraint: The associated LifeStartDate x and BuildingDate y shall be contained into the assigned phase. The BuildingDate shall be greater than the LifeStartDate.

  $\forall p \in$ ST_StructuralArchaeoPart

  $(x = p.\mathtt{LifeStartDate} \wedge y = p.\mathtt{BuildingDate} \wedge$

  $((n.\mathrm{equal}(x) \vee n.\mathrm{before}(x)) \wedge (m.\mathrm{equal}(x) \vee m.\mathrm{after}(x)) \wedge$

  $(\neg y.isUndefined() \implies (n.\mathrm{equal}(y) \vee n.\mathrm{before}(y))\wedge$

  $(m.\mathrm{equal}(y) \vee m.\mathrm{after}(y)) \wedge (y.\mathrm{equal}(x) \vee y.\mathrm{after}(x))))$

  which is represented by the following additional edges:

$$- \ n \xrightarrow{\langle 0,0,d-a,d-a \rangle [1]} x$$

$$- \ x \xrightarrow{\langle 0,0,d-a,d-a \rangle [1]} m$$

$$- \ n \xrightarrow{\langle 0,0,d-a,d-a \rangle [1]} y$$

$$- \ y \xrightarrow{\langle 0,0,d-a,d-a \rangle [1]} x$$

$$- \ x \xrightarrow{\langle 0,0,+\infty,+\infty \rangle [1]} y$$

- ST_ReusedArchaeoPart
  Constraint: The associated `LifeStartDate` x shall precedes the assigned phase. The `ReuseDate` y shall be contained into the phase and be greater than the `LifeStartDate`.

$$\forall p \in \texttt{ST\_ReusedArchaeoPart}$$
$$((x = p.\texttt{LifeStartDate} \ \wedge \ x.\text{before}(n)) \wedge$$
$$(y = p.\texttt{ReuseDate} \ \wedge \ \neg y.isUndefined() \implies$$
$$((n.\text{before}(y) \vee n.\text{equal}(y)) \wedge$$
$$(m.\text{after}(y) \vee m.\text{equal}(y) \wedge y.\text{after}(x)))$$

which is represented by the following additional edges:

$$- \ x \xrightarrow{\langle 0,0,+\infty,+\infty \rangle [1]} n$$

$$- \ n \xrightarrow{\langle 0,0,d-a,d-a \rangle [1]} y$$

$$- \ y \xrightarrow{\langle 0,0,d-a,d-a \rangle [1]} x$$

$$- \ x \xrightarrow{\langle 0,0,+\infty,+\infty \rangle [1]} y$$

9　Example of Reasoning on a SITAVR Model

The translation of a model to a FTCN allows to answer different interesting questions. In particular, in the archaeological domain two issues can be of particular interest: compute the minimal network (i.e., minimize the constraints and find more precise dates), and check the network consistency in order to aid the archaeologist during the dating process.

This section illustrates an example of reasoning performed on a portion of the SITAR model that allows the identification of some new temporal knowledge. It regards an archaeological object called *Porta Borsari* which is an ancient Roman gate in Verona. This object has been modeled as an `ST_ArchaeoUnit` which is composed of seven archaeological partitions and is characterized by three different phases in its existence.

- Phase A: first foundation as *Porta Iovia* during the Late Republican Time (from 200 B.C. to 27 B.C.)

- Phase B: reconstruction during the Claudian Time (from 41 A.C. to 54 A.C.)

- Phase C: Teodorician changes during the Middle-Ages (from 312 A.C. to 553 A.C.).

The three phases are temporally located using the `era` attribute inside the corresponding nodes: in particular, phase A starts and ends inside the Late Republican Time, phase B starts and ends during the Claudian Time, and phase C starts and ends inside the Middle-Ages.

Archaeological partitions are dated as in the second column of Table 4, and assigned to the phase reported in the third column of the same table. For all partitions, only the `LifeStartDate` has been specified; moreover, the

Table 4: Dating of each partition and associated phase.

Archaeological Partition		LifeStartDate	Phase
P208	Foundation and North Tower	$\langle -110, -100, -1, +9 \rangle [1]$ I B.C. $\pm$ 10 years	A
P263	Structures of the eastern facade	$\langle -60, -50, -45, -35 \rangle [1]$ Middle of I B.C. $\pm$ 10 years	A
P214	Front of the external facade	$\langle 35, 45, 50, 60 \rangle [1]$ Middle of I A.C. $\pm$ 10 years	B
P248	External Foundations	$\langle -9, 1, 100, 110 \rangle [1]$ I A.C. $\pm$ 10 years	B
P275	Internal Foundations	$\langle -10, 1, 50, 100 \rangle [1]$ Middle of I A.C. $\pm$ 5 years	B
P250	Defensive structures	$\langle 401, 450, 500, 500 \rangle [1]$ 2nd middle of V A.C.	C

following temporal relations are known between partitions:

- P208 terminates before P263 starts

- P248 terminates before P214 starts

Accordingly with the transformation rules of the previous section, the first operation to perform is the definition of a common coordinate reference system. The origin of such system is placed to 200 B.C., since it is the earliest date in the model, while the interval is year since all dates have the granularity of at least one year.

In order to simplify the presentation, the resulting network is presented through three portions, each one corresponding to a different phase. The overall network can be obtained by combining the three pieces and by adding an edge from phase A to phase B and an edge from phase B to phase C, both labeled with $\langle 0, 0, +\infty, +\infty \rangle[1]$. These edges represents the precedence relations between phases. Moreover, when not specified, α is assumed equal to 1, while the constraint $\langle 0, 0, +\infty, +\infty \rangle[1]$ is usually omitted from the arcs for not cluttering the diagram.

Fig. 17 illustrates the subnetwork related to phase A: node s represents the starting point, nodes A_s and A_e represent the start and end points of the phase respectively, while nodes P263 and P208 represent the `LifeStartDate` of the corresponding archaeological partitions. This portion of FTCN allows to compute some derived constraints for the nodes based on the declared one, using the formula in Def. 5.19: $\pi'_{ij}(x) = \pi_{ij} \otimes_a (\pi_{ik} \circ \pi_{kj}(x))$.
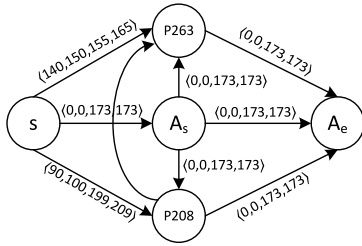

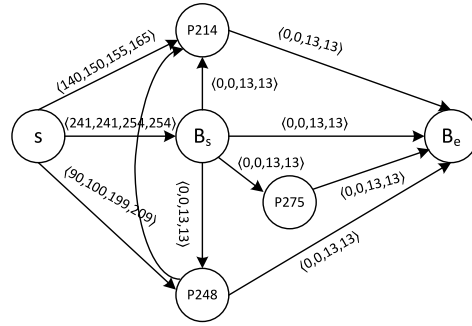Figure 17: Portion of FTCN related to phase A.

Figure 18: Portion of FTCN related to phase B.

In particular, a more precise relation can be derived between partition P208 and partition P263, which is initially represented simply as an edge labeled with the constraint $\langle 0, 0, +\infty, +\infty \rangle$. In particular, by assuming $i =$ P208, $k = s$ and $j =$ P263, the following new constraint π'_{ij} can be derived between P208 and P263:

$$
\begin{aligned}
\pi'_{ij} &= \pi_{ij} \otimes_a (\pi_{ik} \circ \pi_{kj}) \\
&= \pi_{ij} \otimes_a (\pi_{ki}^{-1} \circ \pi_{kj}) \\
&= \langle 0, 0, \infty, \infty \rangle \otimes_a (\langle -209, -199, -100, -90 \rangle \circ \langle 140, 150, 155, 165 \rangle) \\
&= \langle 0, 0, \infty, \infty \rangle \otimes_a \langle -69, -49, 55, 75 \rangle \\
&= \langle 0, 0, 55, 75 \rangle
\end{aligned}
$$

From this derivation follows that the distance between P208 and P263 can be from 0 to 75 years, with great possibility until 55. This is consistent

with the observation that P208 is located in I B.C., but it shall precede the partition P263 which is located in the middle of I B.C, namely P263 has to start living at most 75 years the life start of P208.

A similar operation can be performed on the FTCN portion in Fig. 18, where B_s and B_e represents the start and end points of phase B, respectively. The constraint between partition PA-248 and PA-214 can be restricted as follows by considering $i = $ P248, $k = s$ and $j = $ P214:

$$
\begin{aligned}
\pi'_{ij} &= \pi_{ij} \otimes_a (\pi_{ik} \circ \pi_{kj}) \\
&= \pi_{ij} \otimes_a (\pi_{ki}^{-1} \circ \pi_{kj}) \\
&= \langle 0, 0, \infty, \infty \rangle \otimes_a (\langle -209, -199, -100, -90 \rangle \circ \langle 140, 150, 155, 165 \rangle) \\
&= \langle 0, 0, \infty, \infty \rangle \otimes_a \langle -69, -49, 55, 75 \rangle \\
&= \langle 0, 0, 55, 75 \rangle
\end{aligned}
$$

The consideration is similar to the previous one, since P214 happens in the middle of the I A.C. and P248 is generally dated I A.C. but has to finish before P214 start living.

Finally, as regards to phase C whose corresponding sub-network is reported in Fig. 19, the dating of its only partition can determine a restriction of the phase start as follows by considering $i = s$, $k = $ P250 and $j = C_s$:

$$
\begin{aligned}
\pi'_{ij} &= \pi_{ij} \otimes_a (\pi_{ik} \circ \pi_{kj}) \\
&= \pi_{ij} \otimes_a (\pi_{ik} \circ \pi_{kj}^{-1}) \\
&= \langle 512, 512, 753, 753 \rangle \otimes_a (\langle 601, 650, 700, 700 \rangle \circ \langle -241, -241, 0, 0 \rangle) \\
&= \langle 512, 512, 753, 753 \rangle \otimes_a \langle 360, 409, 700, 700 \rangle \\
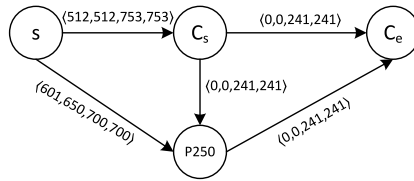&= \langle 512, 512, 700, 700 \rangle
\end{aligned}
$$

Figure 19: Portion of FTCN related to phase C.

Clearly, these are only examples of the derivations that can be obtained by executing the path-consistency algorithm on the overall network and considering all the triangles. However, these examples makes clear the utility of applying known temporal reasoning techniques on archaeological data.

10 Conclusion

This report proposes a framework for representing and managing time dimensions in archaeological data. As regards to the representation task, the applicability of the Standard ISO TC 211 19108 is evaluated by considering a real-world information system, called SITAVR, which has been developed for the archaeological data of Verona.

From this preliminary analysis has emerged that the Standard is unable to represent the inherent vagueness of archaeological data. Therefore, an extension of the Standard concepts has been defined which is based on a fuzzy representation of dates and of ordering relations about time points. Such extension has been successfully applied to the SITAVR case.

Conversely, as concerns to the managing aspect, the main idea is using existing reasoning techniques in order to guide archaeologists during the complex dating process. For this reason, some translation rules have been defined from the proposed extended Standard model to Fuzzy Temporal Constraint Networks (FTCNs). These rules have been applied to a portion of the SITAVR data and some implicit constraints have been derived.

As future work, a tool will be developed for automatically translating an archaeological model into a FTCN using the proposed rules. Such tools will be validated by considering the content of the overall SITAVR information system. Nevertheless, given the nature of archaeological data and the rule of expert knowledge during the interpretation process, the result of an automatic reasoning can provide an invaluable guide during the dating but cannot substitute archeologists in such process.

References

[1] James F. Allen. Maintaining Knowledge About Temporal Intervals. *Communications of the ACM*, 26(11):832–843, 1983.

[2] Silvana Badaloni, Marco Falda, and Massimiliano Giacomin. Integrating Quantitative and Qualitative Fuzzy Temporal Constraints. *AI Communications*, 17(4):187–200, 2004.

[3] Silvana Badaloni and Massimiliano Giacomin. The Algebra IA^{fuz}: A Framework for Qualitative Fuzzy Temporal Reasoning. *Artificial Intelligence*, 170(10):872–908, 2006.

[4] P. Basso, A. Belussi, P. Grossi, and S. Migliorini. Towards the Creation of an Archaeological Urban Information System: Data Modeling of the

Historical Heritage of Verona. In *Proc. of AGILE 2013, Workshop on Integrating 4D, GIS and Cultural Heritage*, pages 1–3, 2013.

[5] Berdien De Roo, Nico Van de Weghe, Jean Bourgeois, and Philippe De Maeyer. The Temporal Dimension in a 4D Archaeological Data Model: Applicability of the GeoInformation Standard. In *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume II-2/W1, pages 111–121, 2013.

[6] Rina Dechter, Itay Meiri, and Judea Pearl. Temporal Constraint Networks. *Artificial Intelligence*, 49(1-3):61–95, 1991.

[7] ISO. *ISO 19108 Geographic Information – Temporal Schema*, 2002.

[8] ISO. *ISO 8601:2004, Data elements and interchange formats – Information interchange – Representation of dates and times*, 2004.

[9] ISO. *ISO 19507:2012, Object Constraint Language (OCL)*, 2012.

[10] M. Katsianis, S. Tsipidis, K. Kotsakis, and A. Kousoulakou. A 3D Digital Workflow for Archaeological Intra-Site Research using GIS. *Journal of Archaeological Science*, 35(3):655–667, 2008.

[11] J. Li, S. Wu, and G. Huang. Handling Temporal Uncertainty in GIS Domain: A Fuzzy Approach. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 34,4/w4, pages 83–87, 2002.

[12] Lluis Vila and Lluis Godo. On Fuzzy Temporal Constraint Networks. *Mathware and Soft Computing*, 3:315–334, 1994.

University of Verona
Department of Computer Science
Strada Le Grazie, 15
I-37134 Verona
Italy

http://www.di.univr.it