Noname manuscript No.
(will be inserted by the editor)

A Spatio-Temporal Framework for Managing Archeological Data

Alberto Belussi · Sara Migliorini

Received: date / Accepted: date

Abstract Space and time are two important characteristics of data in many domains. This is particularly true in the archaeological context where information concerning the discovery location of objects allows one to derive important relations between findings of a specific survey or even of different surveys, and time aspects extend from the excavation time, to the dating of archaeological objects. In recent years, several attempts have been performed to develop a spatio-temporal information system tailored for archaeological data.

The first aim of this paper is to propose a model, called $\mathcal{S}tar$, for representing spatial-temporal data in archaeology. In particular, since in this domain dates are often subjective, estimated and imprecise, $\mathcal{S}tar$ has to incorporate such vague representation by using fuzzy dates and fuzzy relationships among them. Moreover, besides to the topological relations, another kind of spatial relations is particularly useful in archeology: the stratigraphic ones. Therefore, this paper defines a set of rules for deriving temporal knowledge from the topological and stratigraphic relations existing between two findings. Finally, considering the process through which objects are usually manually dated by archeologists, some existing automatic reasoning techniques may be successfully applied to guide such process. For this purpose, the last contribution regards the translation of archaeological temporal data into a Fuzzy Temporal Constraint Network for checking the overall data consistency and reducing the vagueness of some dates based on their relationships with other ones.

Keywords Data Models for Archaeological Data · Fuzzy Temporal Interval · Fuzzy Temporal Constraint Networks · Spatial data · Temporal data

A. Belussi
Strada Le Grazie, 15 - Verona (Italy)
Department of Computer Science - University of Verona
E-mail: alberto.belussi@univr.it

S. Migliorini
Department of Computer Science - University of Verona

1 Introduction

Archaeological data are usually managed through Geographical Information Systems (GISs), since one of their main characteristics is an absolute or relative location on the Earth surface. This type of information, concerning the discovery location of objects, allows one to derive important spatial relations between findings of a specific survey or even of different surveys. This approach is the basis of the well-known stratigraphic analysis, which is one of the main tools adopted by archaeologists to date findings [22]. Together with spatial location, also the temporal dimension is of considerable interest in the archaeological domain, and the two dimensions are often related to each other. This correlation between space and time may be used for deriving new knowledge about the ancient objects of interest. For this reason, some attempts can be found in literature which aim to define a GIS tailored for archaeological data [18], where the spatial dimension is enhanced with temporal aspects.

This paper presents a spatio-temporal archaeological model, called $\mathcal{S}tar$, which has been defined 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, in particular for what regards the Roman period [7,11,10]. SITAVR development 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.

Relative to the spatial dimension, the $\mathcal{S}tar$ model exploits both geometric and topological data types. In particular, point-set topology allows one to describe a spatial scene by specifying the interesting objects and the relations among them avoiding the representation of any details about their location, shape and extent. Geometric and topological primitives can be bound together or not, allowing to represent a completely abstract network of objects, or a partially/completely realized network, where topological primitives are associated with a geometry describing their exact shape, extent and location on the Earth. This qualitative approach for describing a spatial scene can be very useful in many context, such as in archaeology. In particular, in stratigraphic analysis the main information to store is the relative position of findings, while the storage of their exact shape or location is often not so meaningful.

Similarly, as regards to the temporal dimension, the $\mathcal{S}tar$ model 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. Hence, topological structures can be used to explicitly describe relations among time points, even when they cannot be directly derived, since the exact position of time points is not known for sure. These structures found their natural application in the archaeological domain, where precedence relations between objects are frequently better known than their location in time.

Moreover, it has emerged that the time aspects 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 their value. For instance, the construction date of a building can be located with more confidence between 1830 and 1850, plus-minus an additional range of safety of 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 in Section 6.2 of the $\mathcal{S}tar$ model with fuzzy temporal types and relations.

Time knowledge about findings and the relations existing among them are typically used by archaeologists to derive new knowledge or to support the interpretation process. Moreover, a direct connection exists between the spatial relations derived from stratigraphic analysis and the temporal relations among findings, so that additional temporal knowledge can be derived from spatial information. For this reason the paper presents a set of formal rules that allows one to translate spatial-stratigraphic relations into temporal ones. As stated before, such temporal relations will be characterized by a certain level of fuzziness which is determined by the type of the spatial relation existing among the stratigraphic units and an evaluation of their spatial interaction.

In literature many techniques have been proposed for automatically deriving new temporal knowledge from available data. One of this technique is based on the construction of the so called Temporal Constraint Network (TCN) and a fuzzy extension has been developed in [43], 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 [4] for integrating quantitative and qualitative temporal information into a FTCN. The last contribution of this paper is the translation in Section 7 of the information represented with the $\mathcal{S}tar$ model into a FTCN using the previously mentioned rules. Reasoning techniques on FTCN allow one to answer two main issues: (i) checking the network consistency and (ii) computing the minimal network, namely minimizing the number of constraints and finding more precise dates, in order to reduce some vagueness. The proposed solutions of these two issues can be used to guide archeologists in the complex dating and interpretation process. At the end of Section 7, some examples of translation and knowledge derivation are provided.

The overall aim of this paper is to propose a model and a set of formal rules able to exploit and enrich consolidated reasoning techniques, for representing and managing spatio-temporal 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. In particular, it is suitable when some vague time information are known and the available temporal relations can be used to derive new knowledge or reduce the level of uncertainty of the existing one.

2 Related Work

Traditional database systems retain only the latest state of the modeled system or the state at a specific point in time, presenting an up-to-date, but static view of the environment. Conversely, temporal databases provide a uniform and systematic way for dealing with historical data, describing the evolution of information in time. Several query languages [40,39,17] have been proposed to overcome the limitation of SQL with respect to temporal databases. In [40] the authors propose TSQL2, a temporal extension to the SQL-92 Standard query language. Despite many researches proved the usefulness of TSQL2, the project for incorporating some TSQL2 capabilities into the ISO SQL Standard has been canceled in 2001. Subsequently, a draft has been proposed for adding temporal support in SQL Standard, called SQL/Temporal, which includes the support for two temporal dimensions and two semantics for temporal queries [41,42]. Supported temporal dimensions are *valid time*, the time instants or intervals when an information is true in the modeled reality, and *transaction time*, the time interval during which data are current and can be retrieved in the database [25]. An extension of SQL/Temporal has been proposed in [17] where the authors define the T4SQL temporal query language which adds support also for the *availability time*, the time when the database system or user become aware of a fact, and the *event time*, the time when a decision has been taken or an event happened determining the considered fact.

Even if all these time aspects are of great importance in the database field, the archaeological context requires to represent additional and specific time dimensions which are not considered in currently temporal database research. In [26] 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. While database time corresponds to the transaction time described before, the other temporal characteristics can be seen as a specialization of the valid time, each one with a particular meaning that can influence each other producing now knowledge. The $\mathcal{S}tar$ model proposed in this paper includes many of these time categories, in particular: the excavation time, the stratigraphic time (in terms of relative temporal positions between findings), 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 proposing a spatio-temporal model, it is natural to extend existing spatial data models with time, producing a so called *temporal GIS*, as stated in [2]. In particular, two main approaches can be distinguished: systems that model change in time and systems that model time itself. In [28] the authors divide the former in attribute-oriented spatio-temporal databases which track changes in information about spatial entities, and topology-oriented spatio-temporal databases which track changes in positional information about features and their spatial relationships. Similarly, in [36] the authors recognize that changes can affect both spatial and thematic attributes in a GIS and propose a relational method for accessing spatial and temporal topologies.

The model uses only the notion of valid time and a geographical entity goes through a series of historical states of various durations caused by mutations (changes), until it becomes another entity. This kind of representation may reflect the nature of archaeological data and has inspired in some extent the definition of the $\mathcal{S}tar$ model as regards to the definition of the historical phases in finding lifespan.

In [27] the authors recognize the importance of valid and transaction time also within temporal GIS, and a number of model like [44] have been defined which supports both dimensions. In [35] the authors propose a model which extends the relational database of an image database system (referred to as a GIS) to handle both "valid" (transaction) and "effective" (valid) time intervals. It adds new algebraic operations to the standard operations to manipulate the temporal dimensions effectively, demonstrating a successful application of current temporal database research results into a spatial domain.

Standard ISO/TC211 deals with the modeling of geographical information. In particular, ISO Standard 19107 provides concepts for describing the spatial characteristics of geographical information, while ISO Standard 19108 describes the temporal characteristics of geographical information. Many existing GIS systems implement the primitives defined in such Standards. Therefore, they have to be considered in the definition of a spatio-temporal model that aspire to be effectively implemented. A first investigation about the applicability of ISO Standard 19108 for the representation of archaeological data is proposed in [18]. 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. The $\mathcal{S}tar$ model proposed in this paper is based on a set of spatial and temporal primitives compliant with the mentioned Standards and tries to fill such gap by proposing a set of fuzzy temporal primitives.

In [31] 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 regarding 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. The framework proposed in this paper considers time aspects with a meaning enriched w.r.t. the concept of valid time, and that can influence each other producing temporal constraints that are meaningful in a specific context.

A Temporal Constraint Network (TCN) [19] 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) [43]. Moreover, in [4] 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 [3] 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 [5] by the same authors, in order to provide a complete fuzzy interval algebra, called IA^{fuz}. The translation rules proposed in this paper which transform a $\mathcal{S}tar$ model instance into a FTCN, exploit and enrich the ideas proposed in [4].

TCN belongs to the research area known as *temporal reasoning*, which analyzes existing data in order to determine their consistency, answer queries about scenarios satisfying all constraints, and derive missing information. These techniques are particularly useful in the archaeological context, in which incomplete temporal data with some constraints are typically available. Conversely, another temporal research area, known as *temporal data mining* analyzes large amount of temporal information in order to discover existing patterns. Many approaches exist for temporal data mining which are based on various data model and are suitable for different applications. In [32] the authors provide a unified view of such concepts and a guideline for selecting the appropriate method and data model based on the specific purpose. The same authors proposes in [33] a new hierarchical language, called TSKR (Time Series Knowledge Representation), for the formulation of temporal knowledge based on interval time series. This language provides an understandable and compact description of temporal relations and is enriched with efficient algorithms for pattern discovering.

Association rule mining is a classical data mining technique that aims to find interesting associations or correlations among items in a database. In [13], the authors propose a framework based on fuzzy temporal association rules, in order to increase rule expressiveness. In particular, the temporal dimension is useful not only to obtain sequential association rules, periodic or cyclic association rules, calendric association rules, or event-driven association rules, but also to analyze how association rules evolve if datasets are evaluated on different time-slices. Conversely, the use of the fuzzy-set theory allows a linguistic interpretation of the rules and provides means to handle the uncertainty in attribute management. The use of the fuzzy-set theory for mining association rules is also proposed in [30], where the authors discus an approach to discover association rules for fuzzy spatial data. In this case, the authors combine and extend techniques developed in both spatial and fuzzy data mining in order to deal with the uncertainty found in spatial data. Finally, in [12] the authors propose an approach for mining fuzzy association rule from data with both spatial and temporal characteristics. Such approach is based on data cubes and the Apriori algorithm.

Another field of data mining is the clustering analysis which is a technique for breaking data down into related components in such a way that patterns and order become visible. The conventional clustering algorithms, like the k-means algorithm, have difficulties in handling natural data which are often vague and uncertain. For this reason, fuzzy-set theory can be successfully applied also on clustering algorithms. In [37], the authors propose a comparative analysis of two fuzzy clustering algorithms, namely fuzzy c-means and adaptive fuzzy clustering algorithm.

The utility of data mining technique for spatial databases have been investigated [29], but remain unexamined for spatio-temporal information. Adding the temporal element, two kinds of new rules can be discovered: spatio-temporal evolution rules, that describe process or state changes of objects over time [21], and spatio-temporal meta-rules [1] or rules about rules, which describe changes between two rule-sets generated for static snapshot states of the database.

The use of data mining technique can be useful also in the archaeological context in order to discover other kinds of relations between objects that are out the scope of this paper. In particular, the use of fuzzy clustering techniques for archeological data analysis is discussed in [8]. The authors state that for archaeological applications the clustering of data in distinct groups is an important task and the use of fuzzy cluster analysis techniques can aid such activity. In particular, the use of a fuzzy approach can help with real data which usually exhibit an interpretable archaeological structure that does not induce a clear cluster separation.

In general, the use of computational intelligence techniques in archaeology is discussed in [6], where the author analyses if it is possible to automate the archaeological knowledge production, coining the term *computable archaeology*. His conclusion is that bringing artificial intelligence into archaeology introduces new conceptual resources for dealing with the structure and growth of scientific knowledge, thus it provides an invaluable tool for archaeologists in improving their work.

3 Spatio-Temporal Primitives and Relations

This section briefly introduces the spatio-temporal data types and relations that are used by the $\mathcal{S}tar$ model presented in the following sections. In particular, Section 3.1 formally discusses the spatial data types and the topological relations, while Section 3.2 presents the temporal data types and the Allen's temporal relations; finally Section 3.3 defines the concept of stratigraphic or archaeological relations. The spatial and temporal data types are compliant with the ISO/TC211 Standard, which concerns the standardization in the field of digital geographic information, it has been developed in close collaboration with the Open Geospatial Consortium, and as regards to spatial data types has been implemented in currently available GIS systems.

3.1 Spatial Data Types and Relations

The $\mathcal{S}tar$ model uses a set of spatial data types and relations whose definition is compliant with the ISO Standard 19107 [23] (Spatial Schema) which has been implemented in many existing systems, such as PostGIS [34]. Similarly, the set of topological relations existing between them is defined in terms of the well-known 9-intersection model [20].

A *geographical feature* is an abstraction of a real world phenomenon which is associated to a location relative to the Earth suface. As described in the introduction, the location of a geographical feature can be described by means of one or more spatial attributes whose value is given by means of a geometric object or a topological object. Geometry provides quantitative descriptions of spatial characteristics through coordinates and mathematical functions; for instance, geometries describe the shape, dimension, position and orientation of geographical features. Conversely, topology allows one to describe a spatial scene by specifying the interesting objects and the relations among them without representing any details about their locations, namely independently from their geometries. Figure 1 illustrates the hierarchy of spatial types considered by the $\mathcal{S}tar$ model. Under the root type `GM_Object`, two parts of the hierarchy are shown: the one on the left describing atomic values, called *primitives*, which is divided between geometric primitives (`GM_GeometricPrimitive`) and topological primitive (`TP_TopologicalPrimitive`), and the one on the right representing aggregates of topological primitives, called *complexes*, represented by the type `TP_TopologicalComplex`. Notice that the ISO Standard 19107 [23] defines also complexes that are composed of geometric primitives, but they are not considered in the $\mathcal{S}tar$ model.

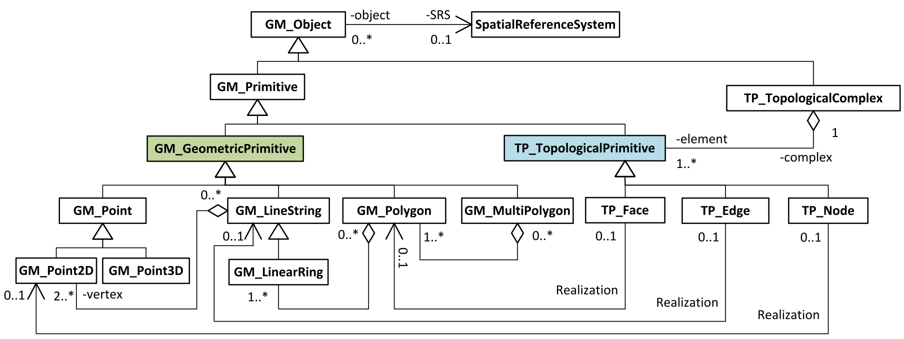


Fig. 1 Hierarchy of geometric and topological spatial types considered in the $\mathcal{S}tar$ model.

The definition of any kind of topological relations among spatial objects is always based on the point-set topological notions of interior, boundary and exterior [20]. For this reason, it is assumed that each geometric and topological type defined in Figure 1 provides a method for retrieving its interior, boundary and exterior. These notions are formally described through the following definitions, considering both the geometric and the topological perspective. Notice that the topological perspective is used here to characterize the behavior of the point-sets representing a geographical feature f in the interaction with other objects, while the topological types of Figure 1 are used for specifying an abstract representation of f.

Definition 1 (Topology) Let X be a set. A *topology* on X is a collection $\mathcal{T}$ of subsets of X that satisfies three conditions: (i) the empty set and X are in $\mathcal{T}$, (ii) $\mathcal{T}$ is closed under arbitrary unions, and (iii) $\mathcal{T}$ is closed under finite intersections. A *topological space* is a set X with a topology $\mathcal{T}$ on it. The sets in a topology $\mathcal{T}$ on X are called *open sets*, and their complements in X are called *closed sets*. □

The following definitions are specified considering the topology $\mathcal{T}$ defined on the reference space X, which is $\mathbb{R}^2$ in the $\mathcal{S}tar$ model.

Definition 2 (Interior) The interior of a geographical feature f, denoted as $I(f)$, is the set of points of f that do not belong to its limit.

Formally, given $Y \subset \mathbb{R}^2$, the interior of Y is the union of all open sets that are contained in Y, i.e., the interior of Y is the largest open set contained in Y. Therefore, a point $y \in \mathbb{R}^2$ is in the interior of Y if, and only if, there is a neighborhood of y contained in Y, i.e., $y \in I(Y)$ if, and only if, there is an open set $U \in \mathcal{T}$ such that $y \in U \subset Y$. □

Definition 3 (Closure) The closure of a geographical feature f, denoted as $C(f)$, is the union of the interior and the boundary of f.

Formally, given $Y \subset \mathbb{R}^2$, the closure of Y is the intersection of all closed sets that contain Y, i.e., the closure of Y is the smallest closed set containing Y. Therefore, y is in the closure of Y if, and only if, every neighborhood of y intersects Y, i.e., $y \in C(Y)$ if and only if $U \cap Y \neq \emptyset$ for every open set $U \in \mathcal{T}$ containing y. □

Definition 4 (Boundary) The boundary of a geographical feature f, denoted as $B(f)$, is the set of primitives that represent its limit. From a geometrical perspective, it is the set of geometries of the next lower dimension that separate the interior from the exterior.

Formally, given $Y \subset \mathbb{R}^2$, the boundary of Y is the intersection of the closure of Y and the closure of the complement of Y, i.e. $B(Y) = C(Y) \cap C(\mathbb{R}^2 \setminus Y)$. Therefore, $y \in B(Y)$ if, and only if, every neighborhood of y intersects both Y and its complement, i.e., $y \in B(Y)$ if, and only if, for all open set $U \in \mathcal{T}$ containing y, $U \cap Y \neq \emptyset$ and $(U \cap (\mathbb{R}^2 \setminus Y)) \neq \emptyset$. □

Definition 5 (Exterior) The exterior of a geographical feature f, denoted as $E(f)$, consists in the difference between the universe $\mathbb{R}^2$ and the closure of f. Formally, given $Y \subset \mathbb{R}^2$, $E(Y) = \mathbb{R}^2 \setminus C(Y)$. □

Given these formal notions, the remainder of this section illustrates the various geometric and topological types defined in Figure 1 and formally defines some topological relations on them. Let us notice that the geometry of an object always depends on the type of coordinate reference system used to define the spatial position. However, in the following we can safely abstract from the particular kind of spatial coordinate reference system and assume that it is known and defined.

Definition 6 (Point) A *point* is a 0-dimensional geometric object which denotes a single location in the coordinate space. It can be represented as a tuple of real numbers representing a 2D/3D coordinate. In the following, the set of possible 2D points in the considered coordinate space is denoted as $\mathcal{P}$, while the set of 3D points is denoted as $\mathcal{P}_3$.

$$\forall p \in \mathcal{P}\ \exists x, y \in \mathbb{R} : p = (x, y)$$
$$\forall p \in \mathcal{P}_3\ \exists x, y, z \in \mathbb{R} : p = (x, y, z)$$

The boundary of a point is the empty set. □

Definition 7 (LineString) A *line-string* is a 1-dimensional object represented by the linear interpolation between a sequence of points, called vertices. Each consecutive pair of vertices defines a *line segment*. In the following, the set of possible line-strings is denoted as $\mathcal{L}$.

$$\forall l \in \mathcal{L}\ \exists v_1, \ldots, v_n \in \mathcal{P} : l = (v_1, \ldots, v_n)$$

The boundary of a line-string consists of its end-points, namely the start and the end vertex. □

Definition 8 (Linear Ring) A *linear ring* is a 1-dimensional geometric object represented by a line-string which is closed and simple. A line-string is closed if its start vertex is equal to its end vertex, while it is simple if it does not pass through the same point twice with the possible exception of the two end points (no self-intersection or self-tangency). In the following, the set of possible linear ring is denoted as $\mathcal{R}$.

$$\forall l \in \mathcal{L} : l \in \mathcal{R} \iff l.isClosed() \land l.isSimple()$$
$$\forall l = (v_1, \ldots, v_n) \in \mathcal{L} : l.isClosed() \iff v_1 = v_n$$
$$\forall l = (v_1, \ldots, v_n) \in \mathcal{L} : l.isSimple() \iff$$
$$(\forall i, j \in [0..1] : (i < j \land M_{[0..1]}(i) = M_{[0..1]}(j)) \implies (i = 0 \land j = 1))$$

where $M_{[0..1]}(x)$ is a function that maps the interval of real numbers $[0..1]$ to the points of the line-string, where $M_{[0..1]}(0) = v_1$ and $M_{[0..1]}(1) = v_n$.

The boundary of a linear ring is the empty set. □

Definition 9 (Polygon) A *polygon* is a 2-dimensional geometric object which is a planar surface defined by one exterior boundary and zero or more interior boundaries (each one defining a hole in the polygon). The exterior and the interior boundaries are represented by linear rings, such that the interior boundaries have an opposite direction w.r.t. the exterior one. In the following, the set of possible polygons is denoted $\mathcal{S}$.

$$\forall s \in \mathcal{S}\ \exists r_e, r_{i_1}, \ldots, r_{i_n} \in \mathcal{R} : s = (r_e, r_{i_1}, \ldots, r_{i_n})$$

□

Definition 10 (MultiPolygon) A *multi-polygon* is a 2-dimensional geometric object composed by one or more polygons, all using coordinates from the same coordinate reference system. The geometric interiors of any two polygons may not intersect, while the boundaries of any two polygons may intersect, at most, at a finite number of points. If two polygons meet along a curve, they could be merged into a single one. In the following, the set of possible multi-polygons are denoted as $\mathcal{M}$.

$$\begin{aligned}
&\forall m \in \mathcal{M}\ \exists s_1, \ldots, s_n \in \mathcal{S} : m = (s_1, \ldots s_n) \wedge \\
&\quad \forall s_i, s_j \in \{s_1, \ldots, s_n\}, i \neq j,\ I(s_i)\ \cap I(s_j) = \emptyset \wedge \\
&\quad \forall s_i, s_j \in \{s_1, \ldots, s_n\}, \forall c_k \in B(s_i), \forall c_h \in B(s_j) \\
&\qquad \exists m \in \mathbb{N} : c_k \cap c_h = \{p_1, \ldots, p_m \mid p_x \in \mathcal{P}, 1 \leq x \leq m\}
\end{aligned}$$

The boundary of a multi-polygon is a set of linear rings corresponding to the boundaries of its Polygon elements. □

Topology deals with the characteristics of geometric shapes that remain invariant if the space is deformed elastically or continuously; for instance, when geographical data is transformed from one coordinate reference system to another one. It is usually applied for describing the connectivity of an n-dimensional graph, in order to convert expensive computational geometry algorithms into combinatorial algorithms and relate geographical features independently from their geometry.

Definition 11 (Spatial topological node) A *spatial topological node* is the 0-dimensional primitive for topology. It can have a realization on the space as a geometric 0-dimensional object, or it can be qualitative described by the spatial relations represented by the edges that start and end in the node. The boundary of a topological node is the empty set. In the following the set of possible spatial topological nodes is denoted as $\mathcal{N}_s$. □

Definition 12 (Spatial topological edge) A *spatial topological edge* is the 1-dimensional primitive for topology. It can have a realization on the space as a geometric 1-dimensional object, or it can be used to describe the relation between its node endpoints. The boundary of an edge is a pair of nodes, the one at the start and the one at the end of the edge. In the following the set of possible spatial topological edges is denoted as $\mathcal{E}_s$.

Definition 13 (Spatial topological face) A *spatial topological face* is the 2-dimensional primitive for topology. It can have a realization on the space as a geometric 2-dimensional object, or it can be used to describe relations among its boundary edges. The boundary of a face is a set of edges with appropriate orientation. In the following the set of possible spatial topological faces is denoted as $\mathcal{F}_s$.

Definition 14 (Spatial topological complex) A *spatial topological complex* is a set of connected topological primitives of all kinds up to the dimension of the complex. It can be represented as a graph and allows one to compactly represent relations among objects. In the following the set of possible spatial topological complexes is denoted as $\mathcal{C}_s$.

In the above defined types the topological approach is used for representing the objects and their spatial properties. However, topology can also be used in the definition of the relations among objects, regardless of which form of representation has been chosen for them (geometric, topological or both). Topological relations are one of the most commonly spatial relations defined between geographical features. They can be computed using the set theoretical operations defined on geometric objects or the algebraic operations defined on topological objects. In particular, this paper refers to the topological relations defined, accordingly to the Dimensionally Extended 9-Intersection Model (DE-9IM) [20], using the so called 9-intersection matrix by testing the intersection between the interior, boundary and exterior of two geographical features.

Given two geographical features a and b of any geometric (or topological) type, the DE-9IM matrix is defined as follows, where $I(a)$, $B(a)$ and $E(a)$ represent the interior, boundary and exterior of a, respectively.

$$M(a,b) = \begin{bmatrix} dim(I(a) \cap I(b)) & dim(I(a) \cap B(b)) & dim(I(a) \cap E(b)) \\ dim(B(a) \cap I(b)) & dim(B(a) \cap B(b)) & dim(B(a) \cap E(b)) \\ dim(E(a) \cap I(b)) & dim(E(a) \cap B(b)) & dim(E(a) \cap E(b)) \end{bmatrix}$$

$dim(x)$ is the maximum dimension of the geometric (or topological) primitives of x, and its possible values are $\{F, 0, 1, 2, T, N\}$ where the value F is returned when x is the empty set, while the value T and N are used to define patterns on the matrix. In such case, when the value T is used in the pattern, it represents any value in the set $\{0, 1, 2\}$, while N means any value in the set $\{F, 0, 1, 2\}$.

Table 1 Topological relation between two features a and b. $I(x)$, $B(x)$ and $E(x)$ represent the interior, boundary and exterior of the feature x, respectively. The relations that can be applied to a polygon (set $\mathcal{S}$) can also be applied to a multi-polygon (set $\mathcal{M}$).

Relation	**Definition**	**Types**	**M(a,b)**
a equals b	$a \subseteq b \wedge b \subseteq a$	any	[TFF FTF FFT]
a disjoint b	$a \cap b = \emptyset$	any	[FFN FFN NNN]
a touch b	$(I(a) \cap I(b) = \emptyset \wedge$ $(a \cap b) \neq \emptyset$	($\mathcal{S}$,$\mathcal{S}$), ($\mathcal{L}$,$\mathcal{L}$), ($\mathcal{L}$,$\mathcal{S}$), ($\mathcal{P}$,$\mathcal{S}$), ($\mathcal{P}$,$\mathcal{L}$)	[FTN NNN NNN] [FNN TNN NNN] [FNN NTN NNN]
a crosses b	$dim(I(a) \cap I(b) \neq \emptyset) <$ $\max(dim(I(a), dim(I(b)) \wedge$ $(a \cap b \neq a) \wedge (a \cap b \neq b)$	($\mathcal{P}$,$\mathcal{L}$), ($\mathcal{P}$,$\mathcal{S}$), ($\mathcal{L}$,$\mathcal{S}$)	[TNT NNN NNN]
		($\mathcal{L}$,$\mathcal{L}$)	[0NN NNN NNN]
a within b	$(a \cap b = a) \wedge$ $(I(a) \cap E(b) = \emptyset)$	any	[TNF NNF NNN]
a overlaps b	$(dim(I(a) = dim(I(b) =$ $dim(I(a) \cap dim(I(b)) \wedge$ $(a \cap b \neq a) \wedge (a \cap b \neq b)$	($\mathcal{P}$,$\mathcal{P}$), ($\mathcal{S}$,$\mathcal{S}$)	[TNT NNN TNN]
		($\mathcal{L}$,$\mathcal{L}$)	[1NT NNN TNN]
a contains b	b within a		

Based on the DE-9IM a set of seven named topological relations have been defined in [16, 14, 15], that are reported in Table 1: {*Equals* (EQ), *Disjoint* (DT), *Touches* (TC), *Crosses* (CR), *Within* (IN), *Contains* (CT), *Overlaps* (OV)}. In the table, the second column provides the point-set definition of the relation, the third column identifies the pair of object types between which the relation can be defined, while the last column contains the corresponding matrix pattern. Notice that, more than one matrix configuration is merged in the same relation.

3.2 Temporal Data Types and Relations

Similarly to the previous section, this section describes the temporal data types and relations used by the $\mathcal{S}tar$ model. In particular, the concepts of geometry and topology, which are typical of the spatial domain, are also applied in the temporal one to describe a temporal object not only in terms of the position it occupies in a temporal reference system, but also in terms of ordering relations.

Figure 2 gives an overall picture of the geometric and topological primitives used in the temporal context. The figure also shows that the hierarchy of the temporal data type is very similar to the corresponding hierarchy of the spatial ones (see Figure 1), in particular for the top levels of the tree.

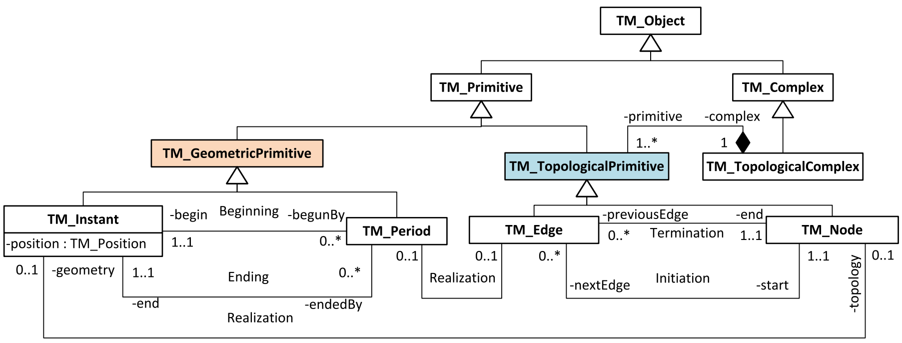


Fig. 2 Hierarchy of geometric and topological temporal types considered in the $\mathcal{S}tar$ model.

Definition 15 (Instant) A *temporal instant* is a 0-dimensional temporal object that identifies a single temporal position on the time axis. Each temporal position can be defined w.r.t. a temporal reference system. In the following, the set of temporal instants is denoted as $\mathcal{I}$.

Three kinds of temporal reference system are considered in this paper: the calendar, the coordinate reference system and the ordinal reference system.

Definition 16 (Calendar) A *calendar* is a discrete temporal reference system that provides a basis for defining temporal positions with a resolution up to one day. A temporal position inside a calendar is called *calDate* and is identified by a tuple of integers $(yyyy, mm, dd)$ whose elements represents the year, month and day, respectively. □

Notice that dates can be defined using different granularities by omitting some parts of the tuple. For instance, the tuple $(2016, 10, 01)$ identifies a date with day granularity, while $(2016, 10, -)$ identifies a date with month granularity.

Definition 17 (Coordinate Reference System) A *coordinate reference system* 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. □

This kind of reference system eases the computation of the distances between temporal primitives and the description of temporal operations.

Definition 18 (Ordinal Reference System) An *ordinal reference system* is a temporal coordinate system based on an ordinal scale. In its simplest form, it is an ordered series of events, having a duration on time axis. □

An ordinal reference system 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 eras. 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. Figure 3 shows the data types used for representing temporal reference systems and how they are involved in the specification of a temporal position.

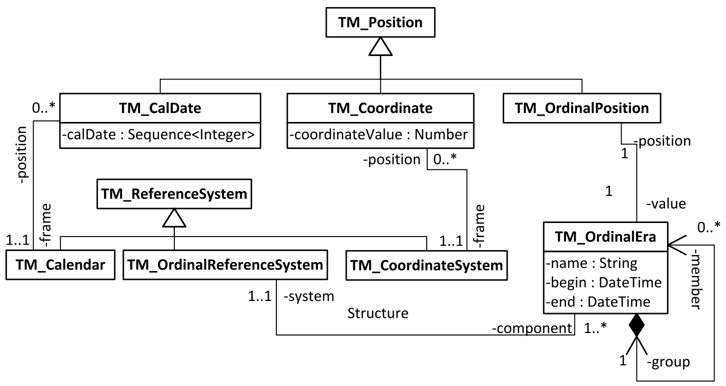


Fig. 3 Definition of a temporal position with respect to a particular reference system.

Definition 19 (Period) A *period* is a 1-dimensional temporal object that represents an interval in the time axis identified by a pair of instants that represents its begin and its end, respectively. In the following the set of periods is denoted as $\mathcal{D}$. □

Table 2 Allen's temporal relations between temporal primitives. Given an instant $i \in \mathcal{I}$, i.pos is its temporal position, while given a period $p \in \mathcal{D}$, p.start and p.end are its start and its end, respectively. $\mathcal{I}_n$ is the set of *Initiation* associations and $\mathcal{T}_e$ is the set of *Termination* associations inside a topological complex $\mathcal{C}$. $(a, b) \in \mathcal{I}_n$ stands for $a \in \mathcal{N}_t \wedge b \in \mathcal{E}_t \wedge$ there exists an *Initiation* association between them. A similar definition holds for $(a, b) \in \mathcal{T}_e$.

Relation	**Types**	**Definition**	**Topological Condition**
a before b	$(\mathcal{I}, \mathcal{I})$	a.pos $<$ b.pos	$\exists$ a sequence $S \in \mathcal{C}$ such that: a is earlier than b in $S \wedge$ $(a, b) \notin \mathcal{I}_n \wedge (b, a) \notin \mathcal{I}_n \wedge$ $(a, b) \notin \mathcal{T}_e \wedge (b, a) \notin \mathcal{T}_e$
	$(\mathcal{P}, \mathcal{I})$	a.end.pos $<$ b.pos	
	$(\mathcal{I}, \mathcal{P})$	b.pos $<$ a.beg.pos	
	$(\mathcal{P}, \mathcal{P})$	a.end.pos $<$ b.beg.pos	
a meets b	$(\mathcal{P}, \mathcal{P})$	a.end.pos $=$ b.beg.pos	$\exists n \in \mathcal{N}_t : (a, n) \in \mathcal{T}_e \wedge$ $(n, b) \in \mathcal{I}_n$
a overlaps b	$(\mathcal{P}, \mathcal{P})$	a.beg.pos $<$ b.beg.pos $\wedge$ a.end.pos $>$ b.beg.pos $\wedge$ a.end.pos $<$ b.end.pos	$\exists n_1, n_2, n_3 \in \mathcal{N}_t : (a, n_1) \in \mathcal{T}_e \wedge$ $(n_2, b) \in \mathcal{I}_n \wedge (n_3, b) \in \mathcal{T}_e \wedge$ n_2 precedes $n_1 \wedge n_1$ precedes n_3
a finishes b	$(\mathcal{I}, \mathcal{P})$	a.pos $=$ b.end.pos	$(a, b) \in \mathcal{T}_e$
	$(\mathcal{P}, \mathcal{P})$	a.beg.pos $>$ b.beg.pos $\wedge$ a.end.pos $=$ b.end.pos	
a contains b	$(\mathcal{P}, \mathcal{I})$	a.beg.pos $<$ b.pos $\wedge$ a.end.pos $>$ b.pos	$\exists n_1, n_2, n_3, n_4 \in \mathcal{N}_t$ $(n_1, a) \in \mathcal{I}_n \wedge (n_2, a) \in \mathcal{T}_e \wedge$ $(n_3, b) \in \mathcal{I}_n \wedge (n_4, b) \in \mathcal{T}_e \wedge$ n_1 precedes n_3 $\wedge$ n_4 precedes n_2
	$(\mathcal{P}, \mathcal{P})$	a.beg.pos $<$ b.beg.pos $\wedge$ a.end.pos $>$ b.end.pos	
a starts b	$(\mathcal{I}, \mathcal{P})$	a.pos $=$ b.beg.pos	$(a, b) \in \mathcal{I}_n$
	$(\mathcal{P}, \mathcal{P})$	a.beg.pos $=$ b.beg.pos $\wedge$ a.end.pos $<$ b.end.pos	
a equals b	$(\mathcal{I}, \mathcal{I})$	a.pos $=$ b.pos	a and b are the same primitive
	$(\mathcal{P}, \mathcal{P})$	a.beg.pos $=$ b.beg.pos $\wedge$ a.end.pos $=$ b.end.pos	
a started by b	same as b starts a		
a during b	same as b overlaps a		
a finished by b	same as b finishes a		
a overlapped by b	same as b overlaps a		
a met by b	same as b meets a		
a after b	same as b before a.		

Instant and period are used to specify the geometric characterization of a temporal primitive. Similarly to what happens for the spatial data types, they have a corresponding topological counterpart in the concepts of node and edge described below.

Definition 20 (Temporal topological node) A *temporal topological node* is a 0-dimensional topological primitive in time. It can have a geometric realization on the time axis as an instant, or it can be qualitative described by means of the relations represented by the edges that start and end in the node. In the following the set of possible temporal nodes is denoted as $\mathcal{N}_t$. □

Definition 21 (Temporal topological edge) A *temporal topological edge* is a 1-dimensional topological primitive in time. It can have a geometric realization as a period, or it can simply represent a temporal relation between two nodes and its corresponding period can be only qualitative described by means of its start and end node. Each edge starts and ends in nodes, while a node can also exists without being associated with edges. In the following the set of possible temporal edges is denoted as $\mathcal{E}_t$. □

Definition 22 (Temporal topological complex) A *temporal topological complex* is a set of connected temporal topological primitives (nodes and edges) that can be represented as a graph. Each edge of a topological complex has start and end nodes inside the complex. □

On the temporal primitives described before, it is possible to defined a set of temporal relations. The $\mathcal{S}tar$ model considers the Allen's relations [3] reported in Table 2. In the table, the second column identifies the type of objects between which the relation can be defined, the third column reports its formal definition, while its definition in terms of topological primitives is shown in the fourth column.

3.3 Archaeological Relations

An *archaeological relation* is the position in space, and by implication in time, of an object or context with respect to another [22]. This kind of relationships are originated from stratigraphy, the main idea is that the spatial relationships that can be determined by observing deposit in section from above, represent the chronological order of their creation.

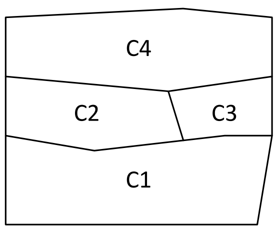


Fig. 4 Example of archaeological relations between contexts.

The principals of stratigraphy are essentially four: *superposition* – the upper units of stratification are younger and the lower are older. *Original horizontality* – archaeological layers deposited in an unconsolidated form will tend towards a horizontal deposition. *Lateral continuity* – any archaeological deposit will be bounded by the edge of the basin of deposition, or will thin down to a feather edge. Therefore, if any edge of the deposit is exposed in a vertical plane view, a part of its original extent must have been removed by excavation or erosion. *Stratigraphic succession* – any given unit of archaeological stratification takes its place in the stratigraphic sequence of a site from its position between the undermost of all units which lie above it and the uppermost of all those units which lie below it and with which it has a physical contact.

The Harris matrix is an effective method used by archeologist to determine the stratigraphic relationships between context. In other words, the position in the matrix determine the position of the contexts in the time sequence. Among all the possible archaeological relations that can be derived using the Harris matrix, this paper concentrates on the following ones:

- *Overlies*: a context overlies another one when the it is vertically above and makes physical contact with the other context. For instance, in Figure 4 context C_4 overlies contexts C_2 and C_3.
- *Above*: a context is above another one when it is vertically above but not necessarily in a physical contact. For instance, in Figure 4 context C_4 is above context C_1. In the same way, we can say that C_4 is also above C_2 and C_3; thus the archaeological relations are not mutually exclusive.
- *Below*: a context is said to be below another one if it is vertically below but not necessarily in physical contact.
- *Contemporary with*: a context is said to be contemporary with another one if they occupy the same vertical position. For instance, in Figure 4 context C_2 is contemporary with C_3.

4 The $\mathcal{S}tar$ Spatio-Temporal Archaeological Model

The Spatio-Temporal ARchaeological model ($\mathcal{S}tar$) presented in this paper has been inspired during the development of an information system, called SITAVR (Sistema Informativo Territoriale Archeologico di Verona), which collects and manages the archaeological data of Verona, a city in northern Italy [7, 11,10]. The need for a specific spatio-temporal model for archeological data comes from the peculiar characteristics of such context. For instance, the need for several time aspects that go further the valid and transaction time, the inherent vagueness of the time locations of objects, the ability to represent knowledge about space and time relations among objects independently from their absolute spatial or temporal positions, the possibility to derive temporal knowledge from spatial relations. In particular, these last two aspects can be useful in many other contexts, such as geology.

The proposed model is based on the set of spatial and temporal data types presented in the previous section. Besides to the definition of the data type hierarchy, a set of constraints are presented which formalizes and explicitly defines some relations between types that have to be satisfied when used in a modeling activity, for instance a containment relation between the spatial attribute of different classes. These constraints are presented exploiting an OCL-link formalism, because the OCL language [24] is the natural way to express formal constraints on UML class diagrams. Moreover, this formulation ease the translation of the constraints into some primitives of a FTCN.

In $\mathcal{S}tar$ model three main objects of interest can be recognized: `ST_InformationSource`, `ST_ArchaeoPart` and `ST_ArchaeoUnit`, which are characterized by some spatial and temporal dimensions described in the following.

An ST_ArchaeoUnit is a complex archaeological entity obtained from an interpretation process performed by the responsible officer. Such an interpretation is done based on some findings (represented by ST_ArchaeoPart instances) retrieved during an excavation process or a bibliographical analysis (represented by ST_InformationSource instances). Therefore, each ST_ArchaeoUnit is connected to one or more constituent ST_ArchaeoParts, each one representing a single result of an excavation or other investigation process. Three kinds of archaeological partitions are considered: it can be a structural element, a mobile element or a reused element. An ST_InformationSource represents the way used to start collecting information about an archaeological partition. Two main kinds of an information source are represented here: excavation, and bibliographical analysis.

The following sections describe the spatial and temporal characteristics of these objects using the primitives illustrated in the Section 3. In particular, Section 4.1 highlights the spatial properties of the archaeological entities, while Section 4.2 their temporal ones. From this preliminary analysis, some weaknesses of the approach are highlighted in Section 6 leading to the definition of an enhanced model which is described in Section 6.2.

4.1 Spatial Primitives

The spatial characteristics of archaeological entities provided by the $\mathcal{S}tar$ model are summarized in Figure 5. An ST_ArchaeoUnit is characterized by the property geometry of type GM_MultiPolygon representing its extent. Each ST_ArchaeoUnit is connected to one or more ST_ArchaeoPart instances (i.e., its components) which have a spatial attribute geometry of type GM_Polygon defining their extent. A containment relation exists between the extent of an archaeological partition and the extent of its corresponding unit: in other words, each GM_Polygon representing the extent of a partition has to be contained inside the GM_MultiPolygon of any of its corresponding units, as formalized in the following constraint. Notice that the extent of an archaeological unit can be larger than the union of its partition extents, because some reconstruction hypothesis can be formulated during the interpretation process.

Constraint 1 (AU-AP Containment) Given an archaeological partition p and its related archaeological units u_i, the polygon representing the extent of p has to be contained inside the extent of any u_i:

$$\forall p \in \texttt{ST\_ArchaeoPart} \; (\; \forall u_i \in \texttt{ST\_ArchaeoUnit} \; (\\ u_i \in p.\texttt{archaeoUnit} \implies p.\texttt{geometry}.\mathit{within}(u_i.\texttt{geometry})))$$

where p.archaeoUnit denotes the set of archaeological unit reachable from the archaeological partition p using the corresponding association. □

Each ST_ArchaeoPart may also be connected with a set of ST_AltimetricPoints which represent meaningful reference points for the object. For

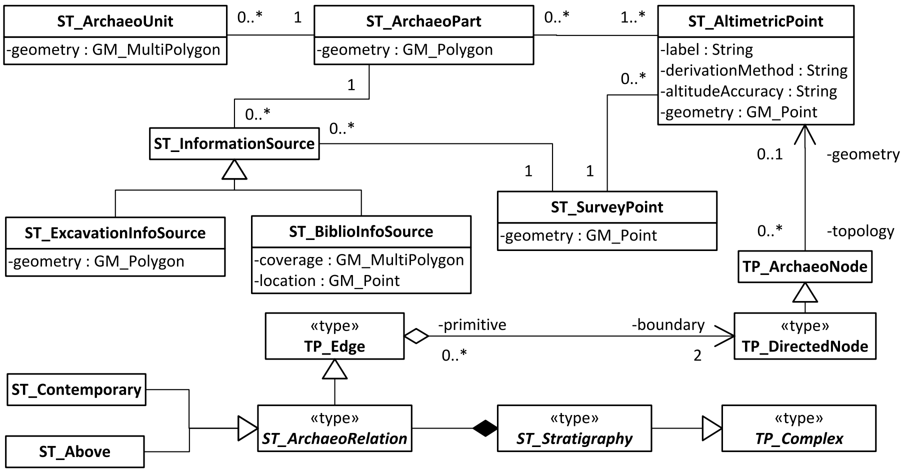


Fig. 5 Spatial components provided by the $\mathcal{S}tar$ model.

the aim of this paper two properties of altimetric points are of major interest: `geometry`, a 3D `GM_Point` representing its position, and `altitudeAccuracy` which defines the degree of reliability of the z value. These two pieces of information, together with the topological relation existing between two archaeological partitions, will be used in the Section 7.3 in order to derive possible archaeological relations between them. In the following, given an archaeological partition/unit x, its minimum and maximum altimetric points are denoted as x_{amin} and x_{amax}, respectively (when a unit is considered, its altimetric points are the union of the altimetric points of all its partitions).

Constraint 2 (AP-AT Containment) Given an archaeological partition p and any of its related altimetric points a_i, the location of a_i has to be contained within the polygon representing the extent of p:

$$\forall p \in \texttt{ST\_ArchaeoPart}\ (\ \forall a \in \texttt{ST\_AltimetriPoint}\ (\\ a \in p.\texttt{altimetricPoint} \implies a.\texttt{geometry}.\mathit{planar}().\mathit{within}(p.\texttt{geometry})))$$

where p.`altimetricPoint` denotes the set of altimetric points reachable from the partition p using the corresponding association, *within* is the topological relation that tests if a geometry is contained inside another one, and *planar* is a method that project in 2D a 3D point by dropping the third coordinate. □

As explained in Section 3.1, topology can be used to represent spatial associations between objects without explicitly define their geometric components. This mechanism can be particularly useful in archeology, in order to represent the stratigraphic relation existing between some findings when their geometry is not known; for instance, because derived from ancient or partial studies. For this purpose a topological complex is defined, called `ST_Stratigraphy`, which is composed of a set of `ST_ArchaeoRelation`s. An archaeological relation is an abstract specialization of `TP_Edge` which can be instantiated as a `ST_Contemporary` or `ST_Above` instance, in order to represent a *contemporary with* or *above* stratigraphic relation, respectively. Notice that the *below* relation can be obtained by using the *above* one and swapping the start and the end nodes. Each archaeological relation connects two topological nodes which are represented by the `ST_ArchaeoPoint` class and can be realized as `ST_AltimetricPoints`, as illustrated in Figure 5.

The spatial component of an `ST_InformationSource` depends upon its type. In particular, two kind of information source are distinguished: `ST_ExcavationInfoSource` and `ST_BiblioInfoSource`. The first one represents an archaeological excavation process which is characterized by an extent of type `GM_Polygon`, while the second one denotes a bibliographic analysis which has two spatial components: an optional `GM_Point` representing the current location of the resource (e.g., an ancient book or cartography), and a `GM_MultiPolygon` representing the territory treated/analysed by the resource. Notice that a containment constraint exists between the extent of an `ST_ExcavationInfoSource` and the extent of its connected `ST_ArchaeoParts`, and similarly between the `coverage` of an `ST_BiblioInfoSource` and the extent of its partitions.

Constraint 3 (EIS-AP Containment) Given an archaeological partition p and its related information source i, the location of p has to be contained into the polygon representing the extent of i, if i is an excavation, or into the polygon representing the coverage of i, if i is a bibliographic source:

$$\begin{aligned}
&\forall p \in \texttt{ST_ArchaeoPart}\ (\ \exists i \in \texttt{ST_InformationSource}\ (\\
&\quad p \in i.\texttt{archaeoPart} \wedge\\
&\quad ((i.\mathit{instanceOf}\,(\texttt{ST_ExcavationInfoSource}) \implies\\
&\qquad a.\texttt{geometry}.\mathit{within}(i.\texttt{geometry})) \vee\\
&\quad\ (i.\mathit{instanceOf}\,(\texttt{ST_BiblioInfoSource}) \implies\\
&\qquad a.\texttt{geometry}.\mathit{within}(i.\texttt{coverage})))))
\end{aligned}$$

where i.`archaeoPart` denotes the set of archaeological partitions reachable from the information source i using the corresponding association, *within* is the topological relation that tests if a geometry is contained inside another one, and *instanceOf* is an OCL function that returns true if the object is an instance of a given class. □

Each `ST_InformationSource` is also connected to a set of `ST_SurveyPoints` which are certified reference points located inside the excavation or the coverage area. They are used to define the altitude of the related `ST_AltimetricPoints`, since its z component is given relative to a particular survey point.

4.2 Temporal Primitives

In the archaeological context time dimension may be specified using different reference systems and different calendars. For this reason, the paper assumes that the reference system and the used calendar are always explicitly declared.

Given an ST_ArchaeoUnit object, a set of possible temporal *phases* of its evolution are identified, then the component ST_ArchaeoPart objects are assigned to one of these phases. This assignment process is one of the fundamental tasks in archaeology [26]. For instance, examples of phases in the existence of an archaeological entity are: installation/foundation, life/use, and renovation/reuse. In *Star* the sequence of phases describing the evolution of

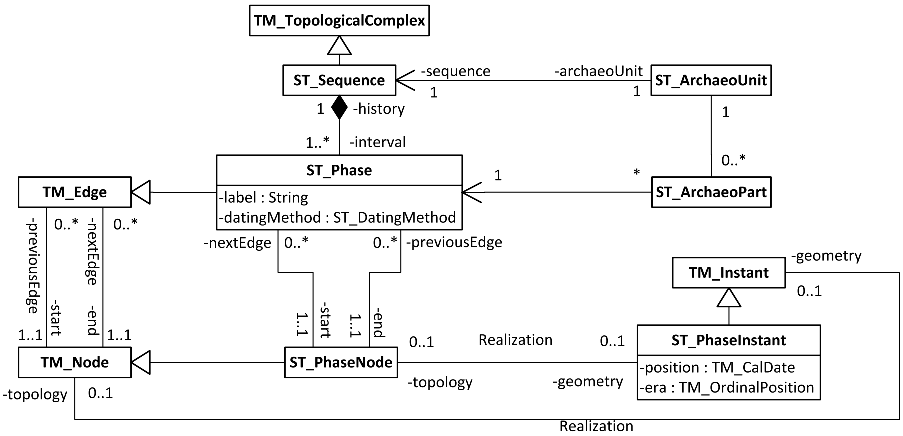


Fig. 6 Representation of the time aspects characterizing an archaeological unit in *Star*.

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 temporal classes of the *Star* model to the types illustrated in Section 3.2, 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 of an ST_ArchaeoUnit object can be modeled using a topological approach. Figure 6 illustrates the result obtained by applying this solution. More specifically, an ST_Sequence can be described as a topological complex which is composed of several 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). *Star* adds two additional properties to each edge: 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 an ST_Phase object with particular nodes (instances of the class ST_PhaseNode specializing TM_Node) which can be realized as a specialization of TM_Instant, called ST_PhaseInstant. Each ST_PhaseInstant has two attributes: a position (inherited from TM_Instant), which is 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 $\mathcal{S}tar$. In case both attributes `position` and `era` are specified, the implicit constraint defined below has to be satisfied.

Constraint 4 (P-E Containment) Given an `ST_PhaseInstant` o, if both its attributes `era` and `position` have been specified, the `position` has to be contained inside the `era`:

$$
\begin{aligned}
&\forall o \in \texttt{ST\_PhaseInstant}(\\
&\quad \neg(o.\texttt{position}.\mathit{isUndefined}() \vee o.\texttt{era}.\mathit{isUndefined}()) \implies\\
&\quad\ ((o.\texttt{era}.\texttt{begin}.\mathit{before}(o.\texttt{position}) \vee o.\texttt{era}.\texttt{begin}.\mathit{equals}(o.\texttt{position}))\wedge\\
&\quad\ \ (o.\texttt{era}.\texttt{end}.\mathit{after}(o.\texttt{position}) \vee o.\texttt{era}.\texttt{end}.\mathit{equals}(o.\texttt{position}))))
\end{aligned}
$$

where `position` and `era` are the two temporal attributes of a phase instant, while `era.begin` and `era.end` are the two endpoints of the temporal interval describing the `era`, and functions *before*, *after* and *equals* test the corresponding temporal relations defined in Table 2.

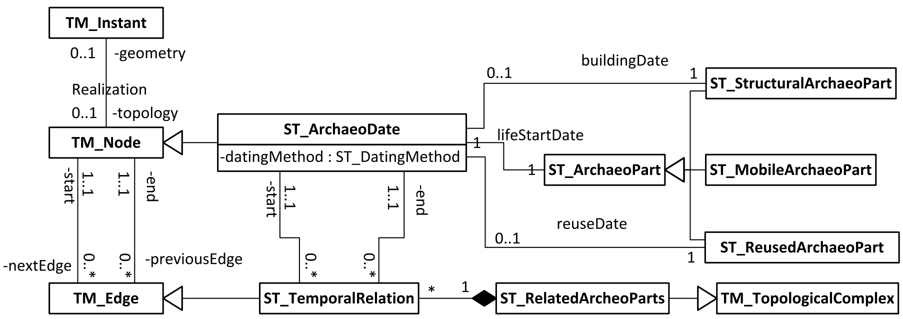


Fig. 7 Representation of the time aspects characterizing an archaeological partition in $\mathcal{S}tar$.

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 represents a structural element, it also has a `buildingDate` role which denotes the date of its construction completion, 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 between the associated phase and the partition dates: for instance, the `lifeStartDate` of a mobile or structural partition shall be contained in the assigned phase, while the `reuseDate` shall precede the phase start node.

Constraint 5 (MAP Constraint) Given an `ST_MobileArchaeoPart` a, its `lifeStartDate` attribute has to be contained inside its assigned phase p.

$$\forall a \in \texttt{ST\_MobileArchaeoPart}\ ($$

$$
\begin{aligned}
&(a.\texttt{phase.start}.\mathit{equals}(a.\texttt{lifeStartDate}) \vee \\
&\quad a.\texttt{phase.start}.\mathit{before}(a.\texttt{lifeStartDate})) \wedge \\
&(a.\texttt{phase.end}.\mathit{equals}(a.\texttt{lifeStartDate}) \vee \\
&\quad a.\texttt{phase.end}.\mathit{after}(a.\texttt{lifeStartDate})))
\end{aligned}
$$

where `lifeStartDate` is the role which identifies the life beginning of a mobile archaeological partition, while `phase` is the phase associated with the object, p.`start` and p.`end` are the two endpoints of the phase p, and functions *before*, *after* and *equals* test the corresponding relations defined in Table 2.

Constraint 6 (SAP Constraint) Given an `ST_StructuralArchaeoPart` a, its `lifeStartDate` attribute has to be contained inside its assigned phase p and its `buildingDate` has to be after or equal to its `lifeStartDate`:

$$
\begin{aligned}
&\forall a \in \texttt{ST\_StructuralArchaeoPart}(\\
&\quad (a.\texttt{phase.start}.\mathit{equals}(a.\texttt{lifeStartDate}) \vee \\
&\quad\quad a.\texttt{phase.start}.\mathit{before}(a.\texttt{lifeStartDate})) \wedge \\
&\quad (a.\texttt{phase.end}.\mathit{equals}(a.\texttt{lifeStartDate}) \vee \\
&\quad\quad a.\texttt{phase.end}.\mathit{after}(a.\texttt{lifeStartDate})) \wedge \\
&\quad (\neg a.\texttt{buildingDate}.\mathit{isUndefined}() \implies \\
&\quad\quad ((a.\texttt{phase.start}.\mathit{equals}(a.\texttt{buildingDate}) \vee \\
&\quad\quad\quad a.\texttt{phase.start}.\mathit{before}(a.\texttt{buildingDate})) \wedge \\
&\quad\quad (a.\texttt{phase.end}.\mathit{equals}(a.\texttt{buildingDate}) \vee \\
&\quad\quad\quad a.\texttt{phase.end}.\mathit{after}(a.\texttt{buildingDate})) \wedge \\
&\quad\quad (a.\texttt{buildingDate}.\mathit{equals}(a.\texttt{lifeStartDate}) \vee \\
&\quad\quad\quad a.\texttt{buildingDate}.\mathit{after}(a.\texttt{lifeStartDate})))))
\end{aligned}
$$

where `lifeStartDate` and `buildingDate` are the roles which identify the life beginning and the building date of a structural object, respectively; while `phase` is the phase associated with the object, p.`start` and p.`end` are the two endpoints of the phase p, and functions *before*, *after* and *equals* test the corresponding temporal relations defined in Table 2.

Constraint 7 (RAP Constraint) Given an `ST_ReusedArchaeoPart` a, its `reuseDate` attribute has to be contained inside its assigned phase p and its `lifeStartDate` has to be before its `reuseDate`:

$$
\begin{aligned}
&\forall a \in \texttt{ST\_ReusedArchaeoPart}\ (\\
&\quad \neg a.\texttt{reuseDate}.\mathit{isUndefined}() \implies \\
&\quad\quad ((a.\texttt{phase.start}.\mathit{before}(a.\texttt{reuseDate}) \vee \\
&\quad\quad\quad a.\texttt{phase.start}.\mathit{equals}(a.\texttt{reuseDate})) \wedge \\
&\quad\quad (a.\texttt{phase.end}.\mathit{after}(a.\texttt{reuseDate}) \vee \\
&\quad\quad\quad a.\texttt{phase.end}.\mathit{equals}(a.\texttt{reuseDate})) \wedge
\end{aligned}
$$

a.reuseDate.*after*(a.lifeStartDate)))

where lifeStartDate and reuseDate are the roles which identify the life beginning and the reuse date of a structural object, respectively; while phase is the phase associated with the object, p.start and p.end are the two endpoints of the phase p, and functions *before*, *after* and *equals* test the corresponding temporal relations defined in Table 2.

The date assigned to an ST_ArchaeoPart object is described in the model by the ST_ArchaeoDate class which is a specialization of the ST_Node one and has consequently an optional realization as an ST_Instant. An additional attribute describing the applied dating method characterizes the ST_ArchaeoDate. The chronology of a partition can also be represented by topological primitives, since a relative order between interactive partitions is better known, than their absolute location. Some edges, called ST_TopologicalRelation, can be placed between nodes representing ST_ArchaeoDates to define temporal relations between related archaeological partition dates. A set of temporal relations relative to some interacting partitions constitutes a topological complex, called ST_RelatedArchaeoParts. In accordance with the model in Section 3.2, the relative positions of two TM_TopologicalPrimitives depend upon the positions they occupy within the sequence of TM_TopologicalPrimitives that make up a TM_TopologicalComplex, as discussed in Table 2 of Section 3.2. The following example illustrates a possible topological structure composed of a set of related archaeological partitions.

Example 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 Figure 8. Dates associated to nodes f_3 and f_4 are realized as 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 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. □

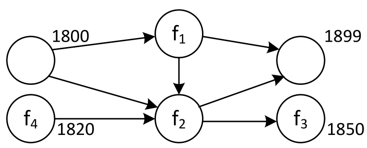


Fig. 8 Example of topological complex representing ordinal temporal relations between chronologies of archaeological partition.

Each ST_ArchaeoPart and each ST_ArchaeoUnit refers to an instance of ST_InformationSource. An ST_InformationSource is characterized by a time dimension that, in accordance to [18], is represented as a geometric primitive, since it is a generally known and documented in some way. 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, as illustrated in Figure 9.

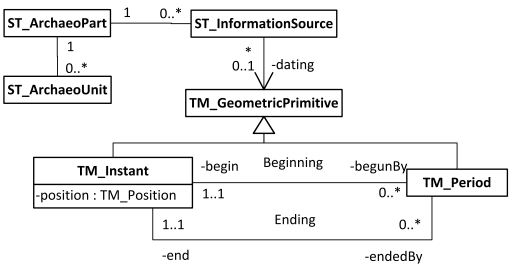


Fig. 9 Representation of the time aspects characterizing an information source in $\mathcal{S}tar$.

5 Representing Vagueness in Time

The $\mathcal{S}tar$ model highlights some weaknesses in the representation of archaeological data, in particular as regards to its temporal dimension. As discussed in the introduction, temporal aspects of archaeological data are typically vague and for this reason many dates are wrongly described as periods instead of instants with the aim to provide a possibility interval for their value. Section 6 will introduce a set of fuzzy temporal data types that allow to overcome such problems. However, in order to understand the rationale behind their definition, this section introduces some basic notions about fuzzy temporal constraint networks and how they can be used to perform reasoning.

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. This paper considers both quantitative (metric) and qualitative (logical) temporal information. In particular, it refers to temporal constraint network and its fuzzy extension for representing metric knowledge, while it considers the Allen's temporal relations for the logical one.

5.1 Fuzzy Temporal Constraint Network

Temporal Constraint Network (TCN) [19] is a formalism for representing temporal knowledge based on *metric* constraints among pairs of time-points. This paper considers only binary constraints, since their expressiveness is satisfactory for many applications.

Definition 23 (Temporal Constraint Network) A *temporal constraint network* $\mathcal{N}$ is a tuple $\langle \mathcal{X}, \mathcal{K} \rangle$, where $\mathcal{X}$ is a set of variables representing time points, and $\mathcal{K}$ is a set of binary constraints on those variables. Variables take values on $\mathbb{R}$, while a constraint k_{ij} restrict the duration of the time elapsed between two temporal variables $x_i, x_j \in \mathcal{X}$ [19]. □

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

Example 2 Let us consider a simple TCN $\mathcal{N}$ composed by a set of temporal variables $\mathcal{X} = \{x, y, z\}$ representing the occurrence of the following events: the beginning of an excavation process, the discovery of a finding f, and the end of the excavation process. Some constraints are defined among such events: the overall duration of the excavation process is between 200 to 230 days, finding f has been discovered between 30 to 40 days after the beginning of the excavation and between 180 to 190 days before the end of the excavation: $\mathcal{K} = \{x \xrightarrow{[200-230]} z, x \xrightarrow{[30-40]} y, y \xrightarrow{[180-190]} z\}$. This network can graphically represented as in Figure 10. □

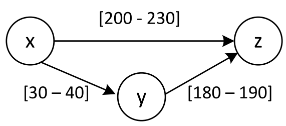


Fig. 10 Graphical representation of the TCN described in Example 2.

However, in the archaeological domain, temporal knowledge is generally characterized by a level of vagueness and dates are usually expressed as periods of great confidence together with a safety additional interval. For instance, the construction date of a building can be expressed as: between 1830-1850 with more confidence, plus or minus 10 years of safety. Fuzzy set theory and probability theory are two related but different ways for modeling uncertainty. In [38] the authors discuss the similarities and differences between these two theories. In particular, they state that probability statements are about the likelihoods of outcomes: an event either occurs or does not, and you can choose on it. This theory is typically used to make predictions and is characterized by only two outcomes: true and false. Conversely, fuzzy set theory was introduced as a mean to model the uncertainty of natural language and is extended to handle the concept of partial truth (or degree of truth). It cannot say clearly whether an event occurs or not and is usually applied for describing happened events. For these reasons, a fuzzy representation of time seams to be the more appropriate solution for the representation of time in archaeology.

A generalization of TCN based on fuzzy sets has been proposed in [43] in order to cope with vagueness in temporal relations. 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 24 (Fuzzy Temporal Constraint Network) A *fuzzy temporal constraint network* $\mathcal{F} = \langle \mathcal{X}, \mathcal{K} \rangle$ consists of a set of variables $\mathcal{X} = \{x_1, \dots x_n\}$ and a set of fuzzy temporal constraints $\mathcal{K} = \{k_{ij} \mid i, j < n\}$ between them. Each constraint k_{ij} is represented as a *possibility distribution function* $\pi_{ij} : \mathbb{R} \to [0, 1]$ that restrict the possible values for the temporal distance $x_j - x_i$ [43]. □

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

This paper considers only trapezoidal distributions since they are sufficiently expressive in practical contexts, while computationally less expensive during the reasoning. They can be represented as a 4-tuple $\langle a, b, c, d \rangle$, where the intervals $[b, c]$ and $[a, d]$ represent the core and the support of the fuzzy set, respectively. In [4] 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 paper assumes that the initial knowledge produced by archeologists is always represented by a trapeze with height equal to one. However, during the reasoning the conjunction of the given constraints can produce trapezes with an height less than one; 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 25 (Fuzzy Trapezoidal Constraint) Given two variables x_i and x_j, a *fuzzy trapezoidal temporal constraint* $k_{ij} = \{T_1, \dots, T_m\}$ 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 [4]. □

The components of a trapeze T_k take values as follows: $a_k, b_k \in \mathbb{R} \cup \{-\infty\}$, $c_k, d_k \in \mathbb{R} \cup \{+\infty\}$, $\alpha_k \in [0, 1]$, $supp(\pi_{T_k}) = \{x : \pi_{T_k}(x) > 0\} = [a_k, d_k]$, and $core(\pi_{T_k}) = \{x : \pi_{T_k}(x) = 1\} = [b_k, c_k]$. Moreover, this paper considers only well-formed trapeze [4]: a trapeze $T = \langle a, b, c, d \rangle$ is *well-formed* if $a \leq b \leq c \leq d$. Given a fuzzy set F, the term support denotes the set of elements with a possibility greater than zero, while the term core denotes the set of elements with a possibility equal to 1 or the other maximum value α_k.

Example 3 Let us consider the situation described in Example 2 and suppose that the constraints defined in $\mathcal{K}$ are characterized by some degree of vagueness as follows: the overall duration of the excavation process is between 200 to 230 days ± 30 safety days, finding f has been discovered between 30 to 40 days ± 7 safety days after the beginning of the excavation, and between 180 to 190 days

$\pm$ 7 safety days before the end of the excavation: $\mathcal{K} = \{x \xrightarrow{\langle 170,200,230,260\rangle[1]} z, x \xrightarrow{\langle 23,30,40,47\rangle[1]} y, y \xrightarrow{\langle 173,180,190,197\rangle[1]} z\}$. This network can graphically represented as in Figure 11. □

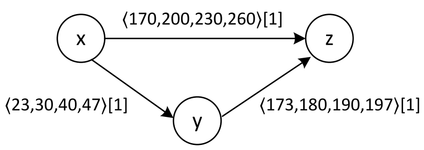


Fig. 11 Graphical representation of the FTCN described in Example 3.

The semantics of a constraint $k_{ij} = \{T_1, \ldots, T_m\}$ 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 26 (Trapezoid Possibility Distribution Function) The *possibility distribution function* of a generic trapeze $T_k \in \mathcal{T}$ can be written as: $\pi_{T_k}(x) = 0$ if $x < a_k \vee x > d_k$, or $\pi_{T_k}(x) = \alpha_k \cdot ((x - a_k)/(b_k - a_k))$ if $a_k \leq x < b_k$, or $\pi_{T_k}(x) = \alpha_k \cdot ((x - d_k)/(c_k - a_k))$ if $c_k < x \leq d_k$, or α_k otherwise [4]. □

Example 4 Let us consider the trapeze $\langle 170, 200, 230, 260\rangle[1]$ defined in the previous example. Its possibility distribution function can be written as follows:

$$\pi_{T_k}(x) = \begin{cases} 0 & \text{if } x < 170 \vee x > 260 \\ \left(\dfrac{x - 170}{200 - 170}\right) & \text{if } 170 \leq x < 200 \\ \left(\dfrac{x - 260}{230 - 170}\right) & \text{if } 230 < x \leq 260 \\ 1 & \text{otherwise} \end{cases}$$

□

Definition 27 (Solution) Let $\mathcal{F} = \langle \mathcal{X}, \mathcal{K}\rangle$ 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 k_{ij} and the degree corresponds to the least satisfied constraint [43]. □

In the case of a FTCN, each solution is characterized by a *degree of satisfaction* reflecting a trade-off among potentially conflicting constraints. Among all possible solutions, the optimal one is the solution that maximizes its degree of satisfaction. The most widely used algorithm for constraint propagation is the *path-consistency algorithm*.

Definition 28 (Path-Consistency Algorithm) Given three variables x_i, x_k and x_j of a FTCN $\mathcal{F}$ and a local instantiation $x_i = d_i$, $x_j = d_j$, a new constraint between x_i and x_j can be induced from pre-existing constraints by the path consistency algorithm as follows: $\pi_{ij} \otimes (\pi_{ik} \circ \pi_{kj})(x)$, where $(\pi_{ik} \circ \pi_{kj})$ is the composition (addition between fuzzy sets) of the constraints between $x_i - x_k$ and $x_k - x_j$, while π_{ij} is the existing constraints between $x_i - x_j$. □

In order to determine the result of the previous definition, it is necessary to define the required operations. More specifically, it is necessary to specialize some operations on fuzzy sets to operations on trapezoids, since not all necessary operations are closed with respect to the trapezoidal form [4]. Notice that these operations can be defined in terms of their constituent trapezoids, since they distribute over disjunction. The specialization of inversion (T_k^{-1}), composition ($T_1 \circ T_2$) and conjunction ($T_1 \otimes_a T_2$) can be found in [9].

Example 5 Let us consider the FTCN in Example 3, the following trapezoid constraints are defined:

$$T_1 = x \xrightarrow{\langle 170,200,230,260\rangle[1]} z$$

$$T_2 = x \xrightarrow{\langle 23,30,40,47\rangle[1]} y$$

$$T_3 = y \xrightarrow{\langle 173,180,190,197\rangle[1]} z$$

from T_2 and T_3 an additional constraint can be derived between x and z using the composition operation: $T_{23} = T_2 \circ T_3 = \langle 196, 210, 230, 244\rangle[1]$. This constraint can be combined with the other constraint T_1 between x and y using the conjunction operation: $T_1 \otimes T_{23} = \langle 196, 200, 230, 244\rangle[1]$. As a result, the combination of the specified constraints produces a more strict possible duration between events x and z, in particular the duration of the safety interval has been restricted from ± 30 days to 4 days for the beginning date and 14 for the ending date. □

5.2 Fuzzy Qualitative Temporal Constraints

Qualitative temporal constraints can be represented using the Allen's Interval Algebra [3]. An extension of this model that integrates the ideas of flexibility and vagueness has been presented in [4, 5] and is called IA^{fuz} algebra.

Definition 29 (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: $I_i(rel_1, \ldots rel_m)I_j$ where each rel_k is one of the 13 mutually exclusive atomic relations: *before*, *after*, *meets*, *metBy*, *overlaps*, *overlapedBy*, *finishes*, *finishedBy*, *contains*, *during*, *starts*, *startedBy*, *equals* [4]. □

In order to integrate the concept of vagueness 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 30 (Fuzzy Qualitative Constraint) Given two temporal intervals I_i and I_j, a fuzzy qualitative constraint k_{ij} between them is represented as: $k_{ij} = (rel_1[\alpha_1], \ldots, rel_{13}[\alpha_{13}])$ where $\alpha_k \in [0,1]$ is the preference degree of rel_k [4]. □

Example 6 Let us consider two intervals I_1 and I_2 which represent the number of days before the discovering of two findings f_1 and f_2, respectively. The fuzzy qualitative constraint k_{23} = (*before* [0.8], *after* [0], *meets* [0.4], *metBy* [0], *overlaps* [0], *overlappedBy* [0], *finishes* [0], *finishedBy* [0], *contains* [0], *during* [0], *starts* [0], *startedBy* [0], *equals* [0]) means that, given the available information, the relation between I_1 and I_2 can be *before* with a preference degree of 0.8, and *meets* with a preference degree of 0.4, while all the other relations are considered as not possible. Notice that it is not necessary that the summation of all α values specified inside a constraint is 1, because it does not represent a probability but a preference degree. □

Accordingly with the model of Section 4 qualitative temporal constraints are represented by topological structures in which temporal nodes are connected through edges. More specifically, each edge denotes a precedence relation between time points. The 13 Allen's temporal relations have been originally defined in terms of interval variables, instead of instant variables. Anyway, some of those relations can be applied also in presence of instant variables, as reported in Table 2. In [4] the authors defines a new algebra PA^{fuzz} in order to express qualitative knowledge concerning points. In particular, the following relations are considered: *before*, *equals* and *after*.

Definition 31 (Fuzzy Qualitative Constraint between Points) Given two time-points p_i and p_j a fuzzy qualitative constraint k_{ij} between them is defined as follows: $k_{ij} = (before[\alpha_1], equal[\alpha_2], after[\alpha_3])$, where *before*, *equal* and *after* are the possible qualitative relations, and $\alpha_k \in [0,1]$ [4]. □

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

Definition 32 (Qualitative to Quantitative) Given a qualitative constraint $k = (before[\alpha_1], equal[\alpha_2], after[\alpha_3])$ between two time points, its corresponding quantitative constraint k_m can be computed as follows: if $\alpha_1 > 0$ then $\langle 0, 0, +\infty, +\infty\rangle[\alpha_1] \in k_m$, if $\alpha_2 > 0$ then $\langle 0, 0, 0, 0\rangle[\alpha_2] \in k_m$, and if $\alpha_3 > 0$ then $\langle -\infty, -\infty, 0, 0\rangle[\alpha_3] \in k_m$. □

Example 7 Let us consider two time points which represent the start of two excavation processes between which the following constraint has been defined k = (*before*[0], *equal*[0.5], *after*[0.8]) which means that the beginning of the first excavation precedes the beginning of the second one with a degree of possibility of 0.8, or they begin simultaneously with a degree of possibility of 0.5. These qualitative relations can be translated into the quantitative constraint $k_m = (\langle 0, 0, +\infty, +\infty\rangle[0.8], \langle 0, 0, 0, 0\rangle[0.5])$. □

6 Modeling Vague Time Dimensions in *Star*

This section discusses the main difficulties in representing vague time information using the *Star* model and how it can be extended in order to overcome such limitations. This analysis leads to the definition of a set of new fuzzy data types that can be used to represent vague time aspects and to perform the reasoning proposed in the previous section.

6.1 Fuzzy Modeling of Temporal Positions

The main lack of the *Star* model 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 Section 4, In particular, as stated in Section 5, this paper concentrates on trapezoidal fuzzy distributions, since they are computationally less expensive, while they provide a sufficient representation of the time knowledge usually provided.

As a general idea, each possible `TM_Position` will be extended in order to express a possibility membership function instead of a certain date. The fuzzy extension of the temporal position is illustrated in Fig. 12.

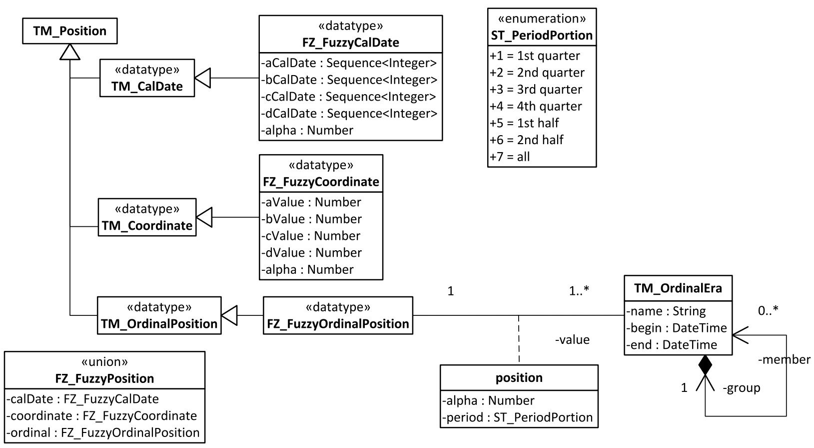


Fig. 12 Fuzzy extension of the temporal positions in the *Star* model.

Each calendar date is represented in a fuzzy form using the `FZ_FuzzyCalDate` datatype which contains a trapezoidal tuple $\langle a, b, c, d\rangle[\alpha]$, where a, b, c, d are sequences of integers representing dates. Similarly, a fuzzy ordinal position inside an ordinal temporal reference system is represented with a specialized class `TM_FuzzyOrdinalPosition`, which has a qualified association with the corresponding `TM_OrdinalEra` enriched 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 as the most possible. Moreover, the cardinality on the era side becomes 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 data types 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.

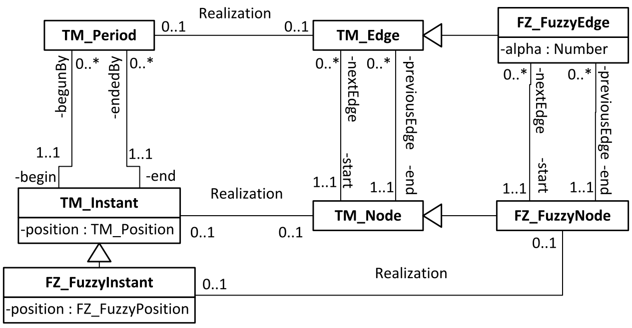


Fig. 13 Fuzzy extension of the topological primitives in the $\mathcal{S}tar$ model.

The last aspect to consider regards the relative ordering between topological primitives inside the same topological complex. In particular, the $\mathcal{S}tar$ model 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 Figure 13. When 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 Figure 13.

6.2 A Fuzzy Extension of the $\mathcal{S}tar$ Model

This section illustrates how the fuzzy data types presented in the previous section can be used for modeling vague time aspects in the $\mathcal{S}tar$ model.

As regards to `ST_ArchaeoUnit`, the extension of their time aspects is illustrated in Figure 14. As explained in Section 4, each archaeological unit is characterized by a sequence of temporal phases describing its evolution. In order, to represent the temporal vagueness in its definition, not only the topological primitives representing the precedence relation among phases have been redefined using the fuzzy types for describing uncertain relations, but also the data type representing the realization of the phase node has been extended.

A phase instant is represented by a `ST_FuzzyPhaseInstant` whose `position` and `era` attributes use the corresponding fuzzy data types.

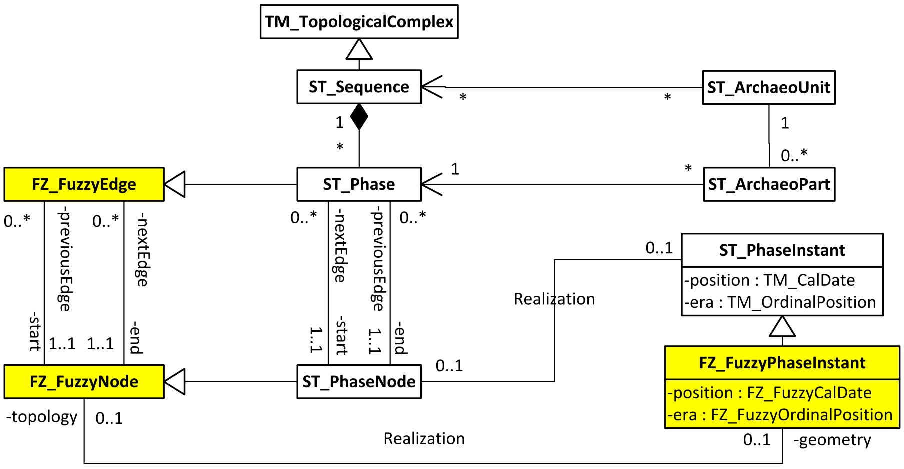


Fig. 14 Fuzzy representation of a `ST_ArchaeoUnit`.

An `ST_ArchaeoPart` has three temporal dimensions: its own dating, the relation with a phase of the corresponding archaeological unit, and 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 Figure 15.

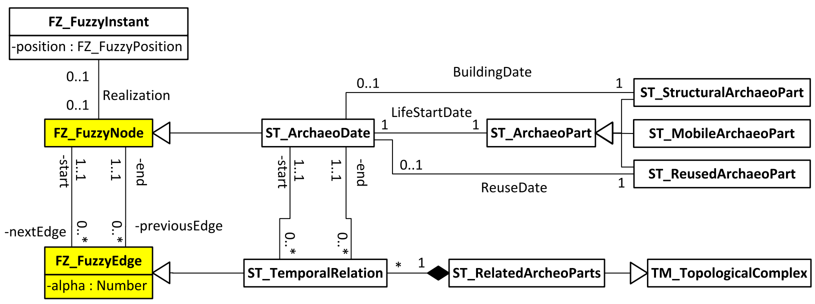


Fig. 15 Fuzzy representation of a `ST_ArchaeoPart`.

7 Translating the $\mathcal{S}tar$ Model to FTCN

This section describes how a $\mathcal{S}tar$ model can be translated into a FTCN in order to derive new temporal knowledge and reduce the existing degree of vagueness. In particular, a set of translation rules are given to encode the structure of the $\mathcal{S}tar$ model, the temporal constraints defined in Section 4, and the qualitative temporal relations given in terms of archaeological relations.

7.1 Translation Rules for the $\mathcal{S}tar$ Classes

To translate the temporal elements introduced in Sections 4 and 6 into a FTCN, it is necessary to initially define a `TM_CoordinateSystem` for transforming all dates into a real number and easing the required comparisons 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 $\mathcal{S}tar$ 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 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: components a and b become the first day of the given month, while c and d become the last day of the given month.
- If g is day and the granularity of x is year: a and b become the first day of the first month of the given year, while 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: a and b become the first month of the given year, c and d become the last month of the given year.
- All the other combinations does not require any transformation. □

Example 8 Let us consider the fuzzy date $x = \langle$1930-03, 1930-03, 1930-12, 1930-12$\rangle[1]$ whose components have a granularity of month, and a desired granularity g of day. It can be translated into a new fuzzy date $x_g = \langle$1930-03-01, 1930-03-01, 1930-12-31, 1930-12-31$\rangle[1]$ where components a and b become the first day of the corresponding month, and c and d become the last day of the corresponding month. The other cases mentioned in Rule 1 can be obtained in a similar way. □

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 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 Figure 16.a, where s is the start node of the network. □

The idea behind such rule is that given the chosen coordinate reference system, each date can be represented as a node x that is connected to the start node of the network by an edge that represents the possible temporal interval between the two nodes. Such temporal instant is expressed w.r.t. the chosen coordinate reference system.

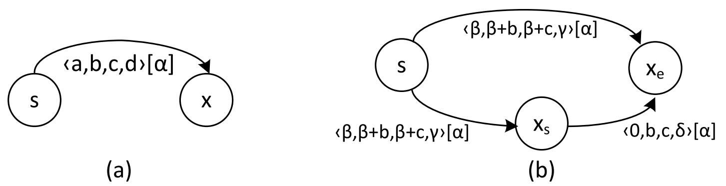


Fig. 16 (a) Translation of a fuzzy calendar date to FTCN. (b) Translation of a fuzzy position inside an ordinal era.

Example 9 Let us consider two fuzzy calender dates $x = \langle 1930\text{-}04, 1930\text{-}06, 1932\text{-}01, 1932\text{-}02\rangle[1]$ and $y = \langle 1930\text{-}05, 1930\text{-}06, 1931\text{-}03, 1931\text{-}04\rangle[1]$ which have to be represented inside a FTCN $\mathcal{F}$ with a granularity g of month. Let us assume that the start node of the network s will represent the date 1930-04 and the temporal values inside a constraint will be given in terms of number of months. The resulting network is illustrated in Figure 17.

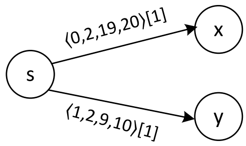


Fig. 17 Example of application of Rule 2 to two calendar dates $x = \langle 1930\text{-}04, 1930\text{-}06, 1932\text{-}01, 1932\text{-}02\rangle[1]$ and $y = \langle 1930\text{-}05, 1930\text{-}06, 1931\text{-}03, 1931\text{-}04\rangle[1]$.

Rule 3 (Coordinate) A `TM_FuzzyCoordinate` x is translated similarly to a `TM_FuzzyCalDate` (Rule 2); however, 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 4 (Ordinal Position) Each `TM_FuzzyOrdinalPosition` related to a `TM_OrdinalEra` x is translated into two nodes x_s and x_e, representing the extremes of the era or of its considered portion. These nodes are connected by an arc labeled as in Table 3 which represents the era duration. Moreover, an arc is added from the start node s to x_s and from s to x_e with a label also defined in Table 3. This translation is illustrated in Figure 16.b. □

Table 3 Translation of the relation between an 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 $\delta = x.\texttt{end} - x.\texttt{begin}$ is the era duration.

Portion	**Arcs** $s \rightarrow x_s$, $s \rightarrow x_e$	**Arc** $x_s \rightarrow x_e$
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	$\langle\beta, \beta, \gamma, \gamma\rangle[1]$	$\langle 0, 0, \delta, \delta\rangle[1]$

In other words, each temporal endpoint of an era is connected to the start node of the network s through an edge which represents the delay between the origin of the chosen reference system and the beginning or the end of the era, respectively. Moreover, the two era endpoints are connected between them by an edge that represents the possible extension of that era.

Example 10 Let us consider the `TM_OrdinalEra` representing the Medieval Ages which comes from 475 a.C. to 1492 a.C., it will be translated as in Figure 18 where nodes x_s and x_e denotes the start and end time points of the era period, while s is the network start node. The two nodes x_s and x_e are connect to s with two edges labeled as $\langle 476, 476, 1492, 1492\rangle[1]$ which represent the interval before the beginning and ending of the era in the chosen reference system, while an edge labeled as $\langle 0, 0, 476, 476\rangle[1]$ connects x_s with x_e to represent the possible era duration. □

Rule 5 (Node) Each not-realized `FZ_FuzzyNode` is translated into a node x and connected with an arc $\langle 0, 0, +\infty, +\infty\rangle[1]$ starting from the network start node s. □

Rule 6 (Edge) Each `FZ_FuzzyEdge` from a `FZ_FuzzyNode` x to a `FZ_FuzzyNode` y is translated into an edge from x to y labeled with the constraint $\langle 0, 0, +\infty, +\infty\rangle[1]$, or $\langle 0, 0, 0, 0\rangle[1]$ depending on the relation existing between them. Notice that an *after* relation from x to y can be translated into a *before* relation from y to x, eliminating the need for the edge $\langle -\infty, -\infty, 0, 0\rangle[1]$. □

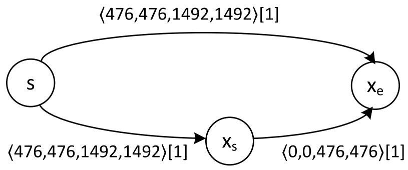


Fig. 18 Example of application of Rule 4 to the `TM_OrdinalEra` representing the Medieval Era which comes from 476 a.C. to 1492 a.C..

Rules 5 and 6 regard the translation of topological constructs which are used to encode qualitative information inside the network. In particular, Rule 5 simply represents the exitance of a temporal node inside the network that happens after the beginning of the chosen coordinate reference system, while Rule 6 encodes a precedence or equality between two temporal instants.

Example 11 Let us consider two archaeological partitions p_1 and p_2 for which it is known that p_1 has been built before p_2. This situation can be modelled using an instance of `ST_TemporalRelation` (i.e., a `FZ_FuzzyTemporalEdge`) between two `ST_ArchaeoDates` (ie., two `FZ_FuzzyNodes`). These two archaeological dates will be translated into two nodes x and y between which an edge is defined with label $\langle 0, 0, +\infty, +\infty\rangle[1]$ that starts in x and ends in y. □

7.2 Translation Rules for the $\mathcal{S}tar$ Constraints

In the $\mathcal{S}tar$ model presented in the Section 4 different implicit constraints have been defined between the spatial and temporal dimensions of object instances. In order to obtain a FTCN which completely represents the content of a $\mathcal{S}tar$ model, it is necessary to incorporate also this information. This section presents a set of rules for translating them into the formalism of FTCNs.

Constraint 4 in Section 4.2 regards the containment between the temporal position associated to a phase and its related era, it can be extended to the `ST_FuzzyPhaseInstant` attributes and translated as follows.

Rule 7 (P-E Containment) Given a `ST_FuzzyPhaseInstant`, its `position` attribute is translated into a node x and connected to the network start node s using Rule 2 or Rule 5, depending on whether it has been realized or not. Similarly, its `era` position is translated into two nodes x_s and x_e and connected to each other and with the network start node as described in Rule 4 or Rule 6, depending on whether it has been realized or not.

Let us assume that the edge between the two nodes x_s and x_e representing the `era` endpoints are connected by an arc $\langle a, b, c, d\rangle[1]$, an arc is added from x_s to y and one from y to x_e with label $\langle 0, 0, c-b, d-a\rangle[1]$ which represent the containment of the phase instant position inside the corresponding era. □

Notice that each `ST_FuzzyPhaseInstant` is always represented by three nodes: two for the era end-points and one for the position, regardless the fact that they are specified in the data or not, since their precedence relation is always valid.

Example 12 Let us assume to have an archaeological partition A which has been assigned to a fuzzy phase whose nodes are realized as two `FZ_FuzzyPhaseInstants` i_1, i_2. For both instants, the reference era is the Medieval Ages described in Example 10. The P-E Containment constraint will be translated as in Figure 19. Notice that besides to the edges between the phase nodes x_s and x_e and the instant nodes i_1 and i_2 which implements Rule 7, an edge has been added between i_1 and i_2 denoting the precedence between the period end-points (see Rule 6). □

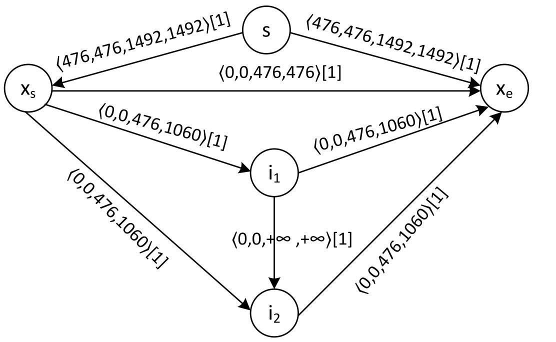


Fig. 19 Example of application of Rule 7 to a phase whose end-points i_1 and i_2 belongs both to the Medieval Ages era represented by nodes x_s and x_e.

Rule 8 (MAP Constraint) Constraint 5 in Section 4.2 regards the containment between the dating of a `ST_MobileArchaeoPart` and its associated phase, and it can be translated as follows. Let us assume that the phase has been translated using Rule 6 and Rule 7, in particular that there exists two nodes n and m which represent the `position` of the start and end phase instants, respectively, and that between these nodes there is an edge $n \xrightarrow{\langle a,b,c,d\rangle[1]} m$ representing the phase duration (i.e., `FZ_FuzzyEdge`). Constraint 5 is represented by two additional edges which connects the position of the `era` instants and the `lifeStartDate` of a mobile archaeological partition:

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

where x is the FTCN node representing the life-start date `FZ_FuzzyNode`. □

The constraint defines an implicit precedence relation between the phase end-points and the life-start date of a mobile archaeological partition: `lifeStartDate` attribute shall be after or equals to the beginning of the phase, and before or equals to the ending of the phase.

Example 13 Let us consider a mobile archaeological partition A whose `lifeStartDate` attribute is represented by the FTCN node x and which is associated with a phase d represented by a start position $p_s = \langle 400, 450, 800, 850\rangle[1]$ and an end position p_e connected with the edge $\langle 0, 50, 100, 150\rangle[1]$ to p_s. This situation can be translated in the FTCN represented in Figure 20 using Rule 8.

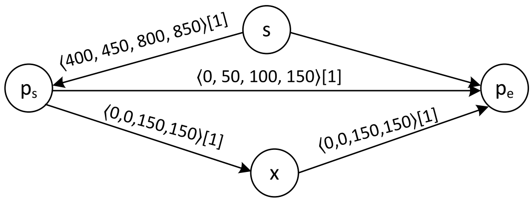


Fig. 20 Example of application of Rule 8 to a mobile archaeological partition A whose life-start date is represented by a FTCN node x which is associated with a phase p.

Rule 9 (SAP Constraint) Constraint 6 in Section 4.2 regards the relation between the dating attributes of a `ST_StructuralArchaeoPart` A and their containment inside the associated phase. The constraint can be translated as follows. Let us assume that the phase has been translated using Rule 6 and Rule 7, in particular that there exists two nodes n and m which represent the `position` of the start and end phase instants, respectively, and that between these nodes there is an edge $n \xrightarrow{\langle a,b,c,d\rangle[1]} m$ representing the phase duration (i.e., `FZ_FuzzyEdge`). The constraint is represented by the additional edges:

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

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

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

where x is the node representing the `lifeStartDate` of p while y is the node representing the `buildingDate` of p.

Let us notice that the two dating attributes shall be contained inside the associated phase; moreover, a precedence relation exists between the `lifeStartDate` and the `buildingDate`.

Example 14 Let us assume to have a structural archaeological partition A which is characterized by a life-start date represented by a node x and building date represented by a node y. Moreover, A is associated to a phase d represented by a start position $p_s = \langle 400, 450, 800, 850\rangle[1]$ and an end position p_e connected with the edge $\langle 0, 50, 100, 150\rangle[1]$ to p_s. This situation can be translated in the FTCN represented in Figure 21 using Rule 9.

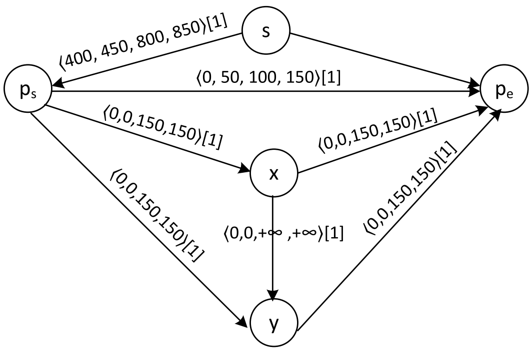


Fig. 21 Example of application of Rule 9 to a structural archaeological partition A whose life-start date is represented by a FTCN node x, the building date is represented by a FTCN node y and which is associated with a phase p.

Rule 10 Constraint 7 in Section 4.2 regards the relations between the dating attributes of a `ST_ReusedArchaeoPart` A and their containment inside the associated phase. The constraint can be translated as follows. Let us assume that the phase has been translated using Rule 6 and Rule 7; in particular, that there exists two nodes n and m which represent the `position` of the start and end phase instants, respectively, and that between these nodes there is an edge $n \xrightarrow{\langle a,b,c,d\rangle[1]} m$ representing the phase duration (i.e., `FZ_FuzzyEdge`). The constraint is represented by the additional edges:

- $n \xrightarrow{\langle 0,0,d-a,d-a\rangle[1]} y$,
- $y \xrightarrow{\langle 0,0,d-a,d-a\rangle[1]} m$, and
- $x \xrightarrow{\langle 0,0,+\infty,+\infty\rangle[1]} y$.

where x is the node representing the `FZ_FuzzyNode` of the life-start date, while y is the node representing the `FZ_FuzzyNode` of the reuse date. Notice that the reuse date can be outside the assigned phase. □

Example 15 Let us assume to have a reused archaeological partition A which is characterized by a life-start date represented by a node x and a reuse date represented by a node y. Moreover, p is associated to a phase d represented by a start position $p_s = \langle 400, 450, 800, 850\rangle[1]$ and an end position p_e connected with the edge $\langle 0, 50, 100, 150\rangle[1]$ to p_s. This situation can be translated into the FTCN represented in Figure 22 using Rule 10.

7.3 Translating Archaeological Relations to Temporal Relations

Given the topological and the archaeological relations existing between two archaeological partitions A and B, a temporal relation can be derived as explained in the following tables. The main idea is to apply the superimposition

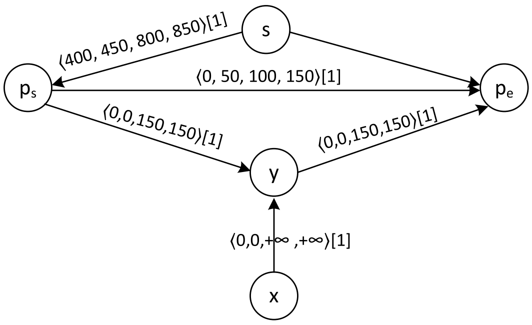


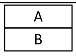
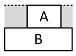
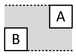
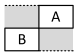
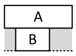
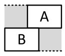
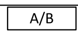
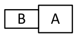
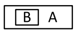
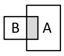
Fig. 22 Example of application of Rule 10 to a reused archaeological partition A whose life-start date is represented by FTCN node x, reuse date is represented by FTCN node y and which is associated with a phase p.

principle for which the upper units of stratification are younger and the lower ones are older. In other words, if an archaeological partition A has a z interval (given by the minimum and maximum elevation of its altimetric points) which is below the z interval of another one B, then A is likely to have a life start date that is before the life start date of B. In particular, the archaeological stratigraphic relation determines the kind of temporal relation, while the topological spatial relation between the 2D projection of the geometries determines the uncertainty of such temporal relation, namely the value of α.

Table 4 shows how the value α is computed considering the different archaeological and topological relation existing between two archaeological partitions A and B. Similarly, Table 5 determines the temporal relation existing between a and b in terms of the arc label connecting two nodes x and y representing their life-start date. As regards to Table 4, the idea is that since the precedence relation is determined from below to above, then the uncertainty is less if the below object fully covers the above one. Therefore, if A is above or overlies B, then the temporal relation is more likely to happens ($\alpha = 1$), while its possibility decreases based on the percentage of overlap, becoming not applicable when the two archaeological partitions become disjoint. In case of a touch, α is set equal to zero, because without an overlap there is no guarantee that the layer containing A is above the layer containing B, it can conversely cuts the other one. The disjoint and touch cases are treated differently in presence of a possible *contemporary with* relation: in such cases, the possibility value α is assumed proportional to the percentage of overlapping of the z coordinate. In other words, the more the two elevation intervals coincide, the more a contemporary relation is possible even without a 2D overlap.

Table 5 determines the kind of the temporal relation existing between two archaeological partitions A and B, given the archaeological relation between them which has been determined using their altimetric point interval. The possible temporal relations between the life-start dates of A and B are those defined in Section 5.2: *before*, *after* and *equals*. Anyway, notice that the archaeological relation *above* and *overlies* have to be both translated with an

Table 4 Computation of the value α of the temporal relation between two archaeological partitions A and B, derived from the archeological and topological relation between them.

Archaeological Relation	Topological Relation	α
A above B or A overlies B	 A equals B	1
	 A within B	1
	 A disjoint B	0
	 A touch B	0
	 A contains B	% of overlap
	 A overlaps B	% of overlap
A below B	same as B above A	
A contemporary with B	 A equals B	1
	 A within B	1
	 A disjoint B	% of z overlap
	 A touch B	% of z overlap
	 A contains B	% of overlap
	 A overlaps B	% of overlap

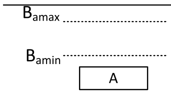
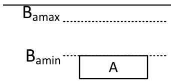
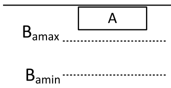
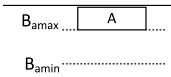
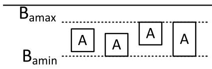
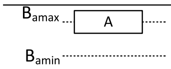
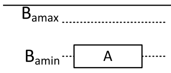
after temporal relation. However, in order to distinguish the two cases, the arc for *above* is labeled as $\langle 1, 1, \infty, \infty \rangle[\alpha]$ where 1 is expressed w.r.t. the minimum temporal granularity in the net, and represents the fact that some minimum time has to elapse between the two dates, while the arc for *overlies* is labeled as $\langle 0, 0, \infty, \infty \rangle[\alpha]$ because the two dates may coincide in the extreme case.

8 Example of Reasoning on a $\mathcal{S}tar$ Model

The translation of a $\mathcal{S}tar$ model into a FTCN allows one 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 archaeologists during the dating process.

This section illustrates an example of reasoning performed on a portion of the SITAR information system 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: 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.), and Phase C – Teodorician changes during the Middle-Age (from 312 A.C. to 553

Table 5 Derivation of the temporal relation between two archaeological partitions A and B, considering the spatial and archaeological relations between them. Let x and y the nodes representing A.`lifeStartDate` and B.`lifeStartDate`, respectively. The last column indicates the label of the arc connecting nodes x and y.

Altimetric Relation		**Archeological Relation**	**Temporal Relation**
	$A.amax < B.amin$	A below B	$x\langle 1,1,\infty,\infty\rangle[\alpha]y$
	$A.amax = B.amin$	B overlies A	$x\langle 0,0,\infty,\infty\rangle[\alpha]y$
	$A.amin > B.amax$	A above B	$y\langle 1,1,\infty,\infty\rangle[\alpha]x$
	$A.amin = B.amax$	A overlies B	$y\langle 0,0,\infty,\infty\rangle[\alpha]x$
	$A.amin \geq B.amin \wedge$ $A.amax \leq B.amax$	A contemp. with B	$x\langle 0,0,0,0\rangle[\alpha]y$
	$A.amin < B.amax \wedge$ $A.amax > B.amax$	A contemp. with B $\vee$ A above B	$y\langle 0,0,\infty,\infty\rangle[\alpha]x$
	$A.amin < B.amin \wedge$ $A.amax > B.amin$	A contemp. with B $\vee$ A below B	$y\langle 0,0,\infty,\infty\rangle[\alpha]x$

A.C.). These 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-Age. The following temporal rela-

Table 6 Dating of each archaeological partition assigned to *Porta Borsari* and definition of the associated phase.

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

tions are also known between partitions: P208 terminates before P263 starts, and P248 terminates before P214 starts. Moreover, there is a partition P220 which has not be dated but has some archaeological relations with partition P214 and P248, whose extent contains the one of P220. In particular, from the altimetric point analysis follows that: P220 *contemporary with* or *below* P214, and P220 *above* P248. Archaeological partitions are dated as in the second column of Table 6, and assigned to the phase reported in the third column of the same table. For all partitions, only the `lifeStartDate` has been specified.

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 described 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.

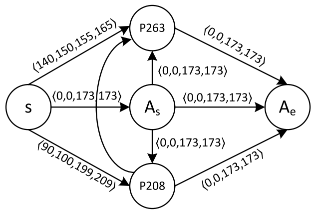


Fig. 23 Portion of FTCN related to Phase A.

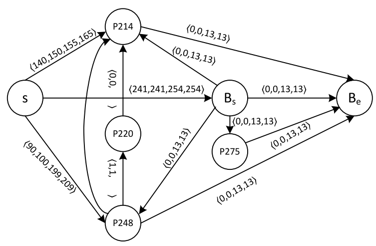


Fig. 24 Portion of FTCN related to Phase B.

Figure 23 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 partition. This portion of FTCN allows to compute some derived constraints for the nodes based on the declared one, using the formula in Definition 28: $\pi'_{ij}(x) = \pi_{ij} \otimes_a (\pi_{ik} \circ \pi_{kj}(x))$. 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 $i = \text{P208}$, $k = s$ and $j = \text{P263}$, the following new constraint π'_{ij} can be derived between P208 and P263: $\pi'_{ij} = \langle 0, 0, \infty, \infty\rangle \otimes_a (\langle -209, -199, -100, -90\rangle \circ \langle 140, 150, 155, 165\rangle) = \langle 0, 0, 55, 75\rangle$. 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 P263 which is located in the middle of I B.C.

A similar operation can be performed on the FTCN portion in Figure 24, where B_s and B_e represents the start and end points of Phase B, respectively. The constraint between partition P248 and P214 can be restricted as follows by considering i = P248, $k = s$ and j = P214: $\pi'_{ij} = \langle 0, 0, \infty, \infty \rangle \otimes_a (\langle -209, -199, -100, -90 \rangle \circ \langle 140, 150, 155, 165 \rangle) = \langle 0, 0, \infty, \infty \rangle \otimes_a \langle -69, -49, 55, 75 \rangle = \langle 0, 0, 55, 75 \rangle$. 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.

As regards to partition P220, starting from its relations with s, P124 and P248, a more precise date can be derived. In particular, let $i = s$, k = P214 and j=P220: $\pi'_{ij} = \langle 0, 0, \infty, \infty \rangle \otimes_a (\langle 140, 150, 155, 165 \rangle \circ \langle 1, 1, \infty, \infty \rangle) = \langle 141, 151, \infty, \infty \rangle$. Given such result and by considering $i = s$, k = P248 and j=P220, the following possibility distribution can be derived $\pi'_{ij} = \langle 141, 151, \infty, \infty \rangle \otimes_a (\langle 90, 100, 199, 209 \rangle \circ \langle -\infty, -\infty, 0, 0 \rangle) = \langle 141, 151, 199, 209 \rangle$.

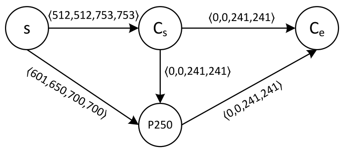


Fig. 25 Portion of FTCN related to Phase C.

Finally, as regards to Phase C, whose corresponding sub-network is reported in Figure 25, the dating of its partition can determine a restriction of the phase start as follows, by considering $i = s$, k = P250 and $j = C_s$: $\pi'_{ij} = \langle 512, 512, 753, 753 \rangle \otimes_a (\langle 601, 650, 700, 700 \rangle \circ \langle -241, -241, 0, 0 \rangle) = \langle 512, 512, 700, 700 \rangle$.

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 make clear the utility of applying existing temporal reasoning techniques on archaeological data.

9 Conclusion

This paper proposes a framework for representing and managing spatial and temporal dimensions in context where a close relation exists between them, such as in the archaeological domain. In particular, a conceptual spatio-temporal archaeological model has been defined. The proposed model can appear unbalanced towards the temporal dimension of archaeological data, since the last scope is to derive new temporal knowledge from spatial information. However, future versions of the model can regard also the fuzzy extension of spatial data types in terms of vague location or extension of findings.

The proposed conceptual spatio-temporal model is called $\mathcal{S}tar$ (Spatio-Temporal ARchaeological model) and has been inspired by a real-world information system, called SITAVR, which collects the archaeological data of Verona. In particular, a fuzzy extension has been proposed which allows to represent vague dates and ordering relations among them. Such extension has been successfully applied to the SITAVR case and a graphical interface is being studied which is suitable for easing the insertion of fuzzy dates and fuzzy temporal relations by archaeologists.

Conversely, as concern 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, a set of formal translation rules have been defined from the proposed conceptual model to Fuzzy Temporal Constraint Networks (FTCNs). Moreover, some derivation rules have been defined which allow one to derive temporal knowledge starting from the spatial and archaeological relations existing between two findings. At the same time, a translation has been provided for some implicit constraints characterizing the model and defined at conceptual level.

As future work, a tool will be developed for automatically translating a conceptual model into a FTCN. Such tool will be validated by considering the content of the overall SITAVR information system and the archaeologist experience. In the context of FTCN and in more generally of TCN, the application of reasoning algorithms is strictly limited by their complexity and in particular by the network dimension. This paper does not consider complexity issues related to the used reasoning algorithms because they have been already studied in literature. However, in a context like the one described in this paper, where spatial and temporal information are closely related, a mechanism can be defined for partitioning the network based on the spatial relations, without reducing or changing the amount and quality of the derived temporal knowledge. Future versions of the framework that will study such partitioning techniques will involved also a study of also the complexity aspect.

References

1. Abraham, T., Roddick, J.: Discovering meta-rules in mining temporal and spatio-temporal data. In: Eighth International Database Workshop, Data Mining, Data Warehousing and Client/Server Databases (IDW'97), pp. 30–41. Springer-Verlag, Hong Kong (1997)
2. Abraham, T., Roddick, J.F.: Survey of Spatio-Temporal Databases. GeoInformatica **3**(1), 61–99 (1998)
3. Allen, J.F.: Maintaining Knowledge About Temporal Intervals. Communications of the ACM **26**(11), 832–843 (1983)
4. Badaloni, S., Falda, M., Giacomin, M.: Integrating Quantitative and Qualitative Fuzzy Temporal Constraints. AI Communications **17**(4), 187–200 (2004)
5. Badaloni, S., Giacomin, M.: The Algebra IA^{fuz}: A Framework for Qualitative Fuzzy Temporal Reasoning. Artificial Intelligence **170**(10), 872–908 (2006)
6. Barceló, J.A.: Computational Intelligence in Archaeology. State of the Art. In: Computer Applications & Qualitative Methods in Archaeology (CAA), Proceedings of the 37th International Conference, pp. 11–21 (2010)
7. Basso, P., Belussi, A., Grossi, P., Migliorini, S.: 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, pp. 1–3 (2013)

8. Baxter, M.J.: Archaeological data analysis and fuzzy clustering. Archaeometry **51**(6), 1035–1054 (2009)
9. Belussi, A., Migliorini, S.: Modeling Time in Archaeological Data: the Verona Case Study. Tech. Rep. RR 93/2014, Department of Computer Science, University of Verona (2014). URL http://www.di.univr.it/report
10. Belussi, A., Migliorini, S., Basso, P., Grossi, P.: The Archaeological Urban Information System of the Historical Heritage of Verona. In: The 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, pp. 1–1 (2015)
11. Belussi, A., Migliorini, S., Grossi, P.: Managing Time Dimension in the Archaeological Urban Information System of the Historical Heritage of Rome and Verona. In: The 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, pp. 235–244 (2014)
12. Calargun, S., Yazici, A.: Fuzzy association rule mining from spatio-temporal data. In: Proceedings of Computational Science and Its Applications ICCSA 2008, *Lecture Notes in Computer Science*, vol. 5072, pp. 631–646. Springer Berlin Heidelberg (2008)
13. Cariñena, P.: Fuzzy temporal association rules: combining temporal and quantitative data to increase rule expressiveness. WIRES Data Mining and Knowledge Discovery **4**(1), 64–70 (2014)
14. Clementini, E., Di Felice, P.: A Comparison of Methods for Representing Topological Relationships. Inf. Sci. Appl. **3**(3), 149–178 (1995)
15. Clementini, E., Di Felice, P.: A Model for Representing Topological Relationships Between Complex Geometric Features in Spatial Databases. Inf. Sci. **90**(1-4), 121–136 (1996)
16. Clementini, E., Felice, P.D., Oosterom, P.v.: A Small Set of Formal Topological Relationships Suitable for End-User Interaction. In: Proceedings of the Third International Symposium on Advances in Spatial Databases, SSD '93, pp. 277–295. Springer-Verlag, London, UK, UK (1993)
17. Combi, C., Montanari, A., Pozzi, G.: The T4Sql Temporal Query Language. In: Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management, CIKM '07, pp. 193–202. ACM, New York, NY, USA (2007)
18. De Roo, B., Van de Weghe, N., Bourgeois, J., De Maeyer, P.: The Temporal Dimension in a 4D Archaeological Data Model: Applicability of the GeoInformation Standard. In: Innovations in 3D Geo-Information Sciences, Lecture Notes in Geoinformation and Cartography, pp. 33–55. Springer International Publishing (2014)
19. Dechter, R., Meiri, I., Pearl, J.: Temporal Constraint Networks. Artificial Intelligence **49**(1-3), 61–95 (1991)
20. Egenhofer, M.J., Franzosa, R.: Point-set topological spatial relations. International Journal of Geographic Information Systems **2**(5), 161–174 (1991)
21. Erwig, M.: Spatio-Temporal Databases: Flexible Querying and Reasoning, chap. Toward Spatio-Temporal Patterns, pp. 29–53. Springer Berlin Heidelberg, Berlin, Heidelberg (2004)
22. Harris, E.C.: Principles of archaeological stratigraphy, 2nd ed. Academic Press (1989)
23. ISO: ISO 19108 Geographic Information – Spatial Schema (2002)
24. ISO: ISO 19507:2012, Object Constraint Language (OCL) (2012)
25. Jensen, C.S., Dyreson, C.E., Böhlen, M., Clifford, J., Elmasri, R., Gadia, S.K., Grandi, F., Hayes, P., Jajodia, S., KÄfer, W., Kline, N., Lorentzos, N., Mitsopoulos, Y., Montanari, A., Nonen, D., Peressi, E., Pernici, B., Roddick, J.F., Sarda, N.L., Scalas, M.R., Segev, A., Snodgrass, R.T., Soo, M.D., Tansel, A., Tiberio, P., Wiederhold, G.: Temporal Databases: Research and Practice, chap. The consensus glossary of temporal database concepts — February 1998 version, pp. 367–405. Springer Berlin Heidelberg, Berlin, Heidelberg (1998)
26. Katsianis, M., Tsipidis, S., Kotsakis, K., Kousoulakou, A.: A 3D Digital Workflow for Archaeological Intra-Site Research using GIS. Journal of Archaeological Science **35**(3), 655–667 (2008)
27. Kemp, Z., Kowalczyk, A.: Incorporating the Temporal Dimension in a GIS. In: M. Worboys (ed.) Innovations in GIS 1, pp. 182–196. Taylor & Francis, London, UK. (1994)
28. Koeppel, I., Ahlmer, S.: Integrating the dimension of Time into AM/FM Systems. In: Proc. of AM/FM XVI Int. Annual Conference (1993)

29. Koperski, K., Adhikary, J., Han, J.: Spatial data mining: Progress and challenges. In: Proc. ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery (DMKD'96), pp. 1–10. ACM Press (1996)
30. Ladner, R., Petry, F.E., Cobb, M.A.: Fuzzy set approaches to spatial data mining of association rules. Transactions in GIS **7**(1), 123–138 (2003)
31. Li, J., Wu, S., Huang, G.: Handling Temporal Uncertainty in GIS Domain: A Fuzzy Approach. In: Int. Arch. of Photogrammetry, Remote Sensing and Spatial Inf. Sciences, vol. 34,4/w4, pp. 83–87 (2002)
32. Mörchen, F.: Unsupervised pattern mining from symbolic temporal data. SIGKDD Explor. Newsl. **9**(1), 41–55 (2007)
33. Mörchen, F., Ultsch, A.: Efficient mining of understandable patterns from multivariate interval time series. Data Min. Knowl. Discov. **15**(2), 181–215 (2007)
34. OSGeo: PostGIS 2.2.3dev Manual. Open Source Geospatial Foundation (2015). http://postgis.net/stuff/postgis-2.2.pdf Accessed March 2016
35. Raafat, H., Xiao, Q., Gauthier, D.A.: An extended relational database for remotely sensed image data management within GIS. IEEE Trans. Geoscience and Remote Sensing **29**(4), 651–655 (1991). DOI 10.1109/36.135827. URL http://dx.doi.org/10.1109/36.135827
36. Raafat, H., Yang, Z., Gauthier, D.A.: Relational spatial topologies for historical geographical information. International Journal of Geographical Information Systems **8**(2), 163–173 (1994). DOI 10.1080/02693799408901992. URL http://dx.doi.org/10.1080/02693799408901992
37. Raju, G., Thomas, B., Tobgay, S., Kumar, T.: Fuzzy clustering methods in data mining: A comparative case analysis. In: Advanced Computer Theory and Engineering, 2008. ICACTE '08. International Conference on, pp. 489–493 (2008)
38. Sanjaa, B., Tsoozol, P.: Fuzzy and probability. In: Int. Forum on Strategic Technology, 2007. (IFOST 2007), pp. 141–143 (2007)
39. Sarda, N.L.: HSQL: A historical query language. In: Temporal Databases, pp. 110–140 (1993)
40. Snodgrass, R.T. (ed.): The TSQL2 Temporal Query Language. Kluwer Academic Publishers (1995)
41. Snodgrass, R.T., Boehlen, M.H., Jensen, C.S., Steiner, A.: Adding transaction time to SQL/temporal (1996). Change proposal, ANSI X3H2-96-502r2, ISO/IEC JTC1/SC21/WG3 DBL MAD-147r2
42. Snodgrass, R.T., Boehlen, M.H., Jensen, C.S., Steiner, A.: Adding valid time to sql/temporal (1996). Change proposal, ANSI X3H2-96-501r2, ISO/IEC JTC1/SC21/WG3 DBL MAD-146r2
43. Vila, L., Godo, L.: On Fuzzy Temporal Constraint Networks. Mathware and Soft Compunting **3**, 315–334 (1994)
44. Worboys, M.F.: A unified model for spatial and temporal information. The Computer Journal **37**(1), 36–34 (1994)