Analyzing Theme, Space, and Time: An Ontology-based Approach

Matthew Perry
LSDIS Lab, University of Georgia
Athens, GA, USA
mperry@cs.uga.edu

Farshad Hakimpour
LSDIS Lab, University of Georgia
Athens, GA, USA
farshad@cs.uga.edu

Amit Sheth
LSDIS Lab, University of Georgia
Athens, GA, USA
amit@cs.uga.edu

ABSTRACT

The W3C's Semantic Web Activity is illustrating the use of semantics for information integration, search, and analysis. However, the majority of the work in this community has focused more on the thematic aspects of information and has paid less attention to its spatial and temporal dimensions. In this paper, we present an integrative ontology-based framework incorporating the thematic, spatial, and temporal dimensions of information. This framework is built around the RDF metadata model. Our ultimate goal is to provide an information system which allows searching and analysis of relationships in any or all of the three dimensions of space, time, and theme. Toward this end, we present an upper-level ontology combining concepts and relationships from both the thematic and spatial dimensions and show how to incorporate temporal semantics into this ontology. We also introduce the notion of a thematic context linking entities of differing dimensions and define a set of query operators built upon these contexts.

Categories and Subject Descriptors
H.2.1 [**Database Management**]: Logical Design – data models

General Terms
Design

Keywords
Ontology, Domain Ontology, Temporal Ontology, Spatial Ontology, Data Modeling, RDF, Spatial Relationships, Temporal, Relationships, Spatio-temporal-thematic Querying

1. INTRODUCTION
It has been said that "an object by itself is intensely uninteresting" [6]. To fundamentally understand any entity, you must examine how it relates to other entities in its world. In other words, relationships are what define semantics (e.g., "relationships at the heart of semantics" [18]). Correspondingly, metadata models defined as part of the W3C's Semantic Web Activity [17] treat relationships as first class objects, and researchers in this community have made progress toward mechanisms for querying complex relationships between resources where an ontology provides the context or domain semantics. Anyanwu and Sheth introduced the concept of semantic associations as complex relationships between resources and defined a set of query operators, ρ, for querying semantic associations in [4]. Semantic associations are defined in terms of connectivity and similarity in RDF graphs. The merits of these complex relationships have been demonstrated in a variety of settings, such as conflict of interest detection in social networks [2] and searching in patent databases [14]. The majority of work in this area has focused almost entirely on thematic relationships between resources, for example the fact that two people deposited money into the same bank account or that two glycopeptides participated in the same biological process.

While thematic metadata can tell us much about how two entities are related, in many domains and applications, we cannot ignore how the entities are related in space and time. The GIS community has put significant effort into the ontological modeling of geospatial relationships and geographic entities and the use of ontologies for search and analysis of geospatial data [1]. However, the power of information systems which integrate ontologies describing thematic aspects of entities with ontologies describing the geospatial and temporal world in which they interact has yet to be fully realized. An information system which captures all of these aspects has enormous potential in many application areas, such as national security, emergency response, and e-learning.

Working towards spatiotemporal and thematic metadata analysis, this paper makes the following contributions: an upper-level ontology which outlines basic classes and relationships for linking the thematic and spatial domains, incorporation of temporal semantics into this ontology, a formalization of spatial and temporal query operators, and a demonstration of the expressiveness of the formalization using example queries based on the scenario of analyzing historical entities and events of World War II. This work is presented in the context of the Resource Description Framework (RDF) metadata model. A unique aspect of our approach is that we do not require the spatial properties of each thematic entity to be explicitly recorded. Instead, we utilize relationships in the thematic domain to indirectly provide spatial properties. This gives the benefit of greater flexibility in the integration of thematic and spatial information, which is necessary for utilization of disparate and incomplete information sources, such as data available on the Web.

2. MODELING THEME, SPACE, AND TIME

This section discusses our approach to modeling theme, space, and time. We present an upper-level ontology defining a general hierarchy of thematic and spatial entity classes and associated relationships connecting these entity classes. We intend for application-specific domain ontologies in the thematic dimension to be integrated into the upper-level ontology through subclassing of appropriate upper-level classes and relationships. Temporal information is integrated into the ontology by labeling relationship instances with their valid times. This ontology-based model is shown in Figure 1.

2.1 Thematic Dimension

Our upper-level thematic ontology consists of a fundamental class hierarchy and a few basic relationships. In developing the class hierarchy, we first follow the approach of Grenon and Smith's Basic Formal Ontology [9] and distinguish between *Continuants* and *Occurrents*. *Continuants* are those entities which persist over time and maintain their identity through change. Examples include a soldier, an aircraft or a city. *Occurrents* represent events and processes; they happen and then no longer exist. Examples are the bombing of a target or the execution of a training exercise. A second division of entities concerns spatial properties. Some *Occurrents* are inherently spatial such as a battle; others are not, such as the assignment of a solider to a division. We therefore explicitly represent *Spatial Occurrents*. *Continuants* also have varying spatial properties. We distinguish a special type of *Continuant* which we refer to as a *Named Place*. *Named Places* are entities which serve as locations for other physical entities and *Spatial Occurrents*. They have very static spatial behavior over time and are distinguished by a strong association with their spatial location. *Named Places* are essentially treated as regions of space in our model, and examples include a city, a zip code, a building, or a lake. In contrast to a *Named Place*, we distinguish another subclass of *Continuant*: *Dynamic Entity*. *Dynamic Entities* are those entities with dynamic spatial behavior whose identities are not as strongly associated with space. Examples include a person or a vehicle. We do not make further philosophical distinctions between these two types of *Continuants* as the final decision depends upon the domain and application.

2.2 Spatial Dimension

In order to search and analyze spatial relationships, we must first model these relationships. There has been much work on the ontological modeling of space. For an excellent survey, see [5]. We are interested in only two-dimensional space and want to support both topological relations and quantitative relations. These are common goals for many geo-ontologies such as the base geospatial ontology described in [12]. The basic classes and relationships we identify mirror those in [11]. The three main entity types are *Spatial Region*, *Coordinate*, and *Coordinate System*. *Spatial Region* has three subclasses: *Centroid*, *Polyline*, and *Polygon*, which are specified with instances of *Coordinate*, and a *Coordinate* is specified in terms of a *Coordinate System*.

We use a handful of important relationships to connect knowledge captured in the spatial ontology with knowledge in the thematic ontology. We define a relationship *occurred_at* which connects *Spatial Occurrent* to *Spatial Region*, and we define a relationship *located_at* which connects *Named Place* to *Spatial Region*. These relationships allow us to associate a thematic concept, such as the city of Berlin or the Battle of the Bulge, with its geospatial properties, which are explicitly and unambiguously specified. Consequently, spatial properties of thematic entities can be derived using the associated *Spatial Regions*.

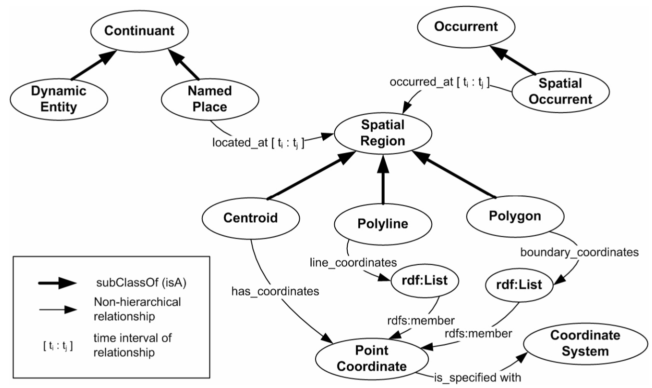

Figure 1. Ontology-based model of space, time, and theme. *Spatial Occurrents* and *Named Places* are directly linked with *Spatial Regions* which record their geographic location. Temporal intervals on relationships denote when the relationship holds (valid time).

2.3 Temporal Dimension

The representation of temporal relationships and temporal properties of information is also a well researched topic. For an overview of the conceptual modeling of time, see [20]. In our work, we are most concerned with exploring temporal relationships over the history of interactions between entities. Therefore, we will adopt a discrete, linearly-ordered time domain, and we will focus on absolute time. The specific relations we are interested in are those of Allen's interval algebra [3] and Ladkin's extension of it over unions of convex time intervals [13] and also temporal distance (e.g. the units of time between two time points).

We incorporate the time dimension into our model by associating time intervals with relationship instances in the ontology. This follows the approach taken by Gutierrez et al. in [10]. The time interval on the relationship denotes the times at which the relationship is valid. For example, consider a soldier assigned to the 1st Armored Division from April 3, 1942 until June 14, 1943 and then assigned to the 3rd Armored Division from June 15, 1943 until October 18, 1943. The relationship connecting the soldier to the 1st Armored Division would be annotated with the closed interval [04:03:1942, 06:14:1943] and the relationship connecting the soldier to the 3rd Armored Division would be annotated with the closed interval [06:15:1943, 10:18:1943]. Any temporal ontology which defines a vocabulary of time units (e.g. Days, Months and Years) can be used to precisely specify the start and end points of time intervals.

3. QUERYING THEME, SPACE, AND TIME

The basic goal of this framework is an extension of thematic analytics which supports search and analysis of spatial and temporal relationships between entities. In this section, we introduce a set of query operators envisioned as an extension of the set of operators, ρ, defined in [4]. The key idea is that non-spatial entities (note that non-spatial entities refer to entities not directly connected to *Spatial Region* in our upper-level ontology) indirectly obtain spatial properties through their relationships with spatial entities, specifically *Named Places* and *Spatial Occurrents*. The nature of the links in these connecting relationships serves as a *context* for the spatial connection. Similarly, entities obtain temporal characteristics indirectly through the temporal properties of their relationships. We define a thematic operator to precisely extract these connecting relationships. The remaining operators fall into two basic groups: those intended to extract spatial and temporal properties for a single entity and those intended to search and analyze spatial and temporal relationships between entities.

We first present the data model on which our query operators are defined. This is followed by a discussion of operators over the thematic dimension, the spatial dimension, and the temporal dimension. Finally, we demonstrate how to combine operators to analyze spatiotemporal aspects of entities.

3.1 Data Model

RDF has been adopted by the W3C as a standard for representing metadata on the Web [16]. Resources in RDF are identified by Uniform Resource Identifiers (URIs) that provide globally-unique and resolvable identifiers for entities on the Web, yielding a decentralized information space. These resources are described through participation in relationships. Relationships in RDF are called *Properties* and are binary relationships connecting resources to other resources or resources to literal values (e.g. Strings or Numbers). These binary relationships are encoded as triples of the form (*Subject*, *Property*, *Object*), which denotes that a resource – the *Subject* – has a *Property* whose value is the *Object*. These triples are referred to as *Statements*. We call a set of triples an *RDF graph*, as RDF data can be represented as a directed, labeled graph with typed edges and nodes. In this model, a directed edge labeled with the *Property* name connects the *Subject* to the *Object*. RDF Schema (RDFS) provides a standard vocabulary for describing the classes and relationships used in RDF statements.

Our spatiotemporal data model incorporates temporal labels on relationship instances, and adopts temporal RDF graphs, defined in [10], to incorporate temporal information into RDF. Given a set of discrete, linearly ordered time points T, a temporal triple is an RDF triple with a temporal label $t \in T$. We use the notation $(s, p, o) : [t]$ to denote a temporal triple. The expression $(s, p, o) : [t_1, t_2]$ is a notation for $\{(s, p, o) : [t] \mid t_1 \leq t \leq t_2\}$. A temporal RDF graph is a set of temporal triples. To deal with incompleteness of web-based data, we assume that statements without timestamps are eternal (always true). A temporal path $p_t = e_1.p_1.e_2.p_2.e_3 \ldots e_{n-1}.p_{n-1}.e_n$ over a temporal RDF graph G is a sequence of URIs in G such that $\forall i$=1..n-1 at least one of $(e_i, p_i, e_{i+1}) : [t]$ or $(e_{i+1}, p_i, e_i) : [t]$ is a temporal triple in G for some $t \in T$ and $\forall i$=1..n, $\forall j$=1..n, if $i \neq j$, then $e_i \neq e_j$ (in other words, a simple, undirected path). We denote the set of temporal triples associated with a temporal path p_t as $TEMPORAL\_TRIPLES(p_t)$. For a temporal path p_t, $TRIPLES(p_t)$ denotes the set $\{(s, p, o) \mid \exists t \in T$ with (s, p, o): $[t] \in TEMPORAL\_TRIPLES(p_t)\}$. For a temporal path p_t, we denote the set $\{t \mid (s, p, o) : [t] \in TEMPORAL\_TRIPLES(p_t)\}$ as $TEMPORAL\_EXTENT(p_t)$.

3.2 Thematic Context

The most basic concept in our approach is what we term a thematic context. Intuitively, a thematic context specifies a type of connection between resources in the thematic dimension of our ontology. For example, connecting two soldiers via membership in a common company could represent one context whereas connecting two soldiers via attendance at the same high school could represent a different context. As mentioned earlier, thematic contexts connecting entities to *Named Places* and *Spatial Occurrents* are of special importance because these connections provide spatial properties for non-spatial entities, for example linking a soldier to the spatial locations of the battles in which he fought.

Here, we define a thematic context for a temporal RDF graph. For a temporal RDF graph G, let $CLASSES(G)$ denote the set of classes in G, let $INSTANCES(G)$ denote the set of instances of all classes $C \in CLASSES(G)$, and let $PROPERTIES(G)$ denote the set of properties in G. For a temporal RDF Graph G, a thematic context tc is an ordered sequence $C_1.P_1.C_2.P_2.C_3 \ldots C_{n-1}.P_{n-1}.C_n$ where $C_i \in CLASSES(G) \cup INSTANCES(G)$ $\forall i$=1..n and $P_i \in PROPERTIES(G)$ $\forall i$=1..n-1. We say a temporal path $p_t = e_1.p_1.e_2.p_2.e_3 \ldots e_{n-1}.p_{n-1}.e_n$ satisfies a thematic context $tc = C_1.P_1.C_2.P_2.C_3 \ldots C_{n-1}.P_{n-1}.C_n$ if $\forall i$, $p_i = P_i$ or p_i is a subproperty of P_i and $e_i \in INSTANCES(G)$ and either $e_i = C_i$, e_i is an instance of C_i, or e_i is an instance of some class C_i' which is a subclass of C_i.

We define one type of thematic query, *ρ-theme*, below:

ρ-theme$(G, tc) \rightarrow \{p_t\}$	
Given:	A temporal RDF graph G, a thematic context tc
Find:	All temporal paths p_t in G which satisfy tc

To connect a *Soldier*, 'John Smith,' to any *Bombings* in which he participated, we could use the *ρ-theme* query in Example 1[1].

ANS ← *ρ-theme* (*G*, 'John Smith'.*on_crew_of*.*Military_Vehicle*.*used_in*.*Bombing*)	(1)

3.3 Spatial Operators

To provide a means to query spatial relationships between resources, we first define operators to extract the spatial properties of a single entity and then define operators to query relationships between these extracted spatial properties.

The first spatial operator we define, *ρ-spatial_extent*, is intended to find the spatial location of a given thematic entity by retrieving the *Spatial Region* associated with it. The query "where were the battles in which the 101st Airborne Division fought?" (see Example 2) represents a typical case for *ρ-spatial_extent*. In our model, we can think of this operator as retrieving the *Spatial Region* connected to the end of a thematic context instance (path). In this case, connected refers to the relationship *occurred_at* or *located_at* connecting a spatial entity (*Named Place* or *Spatial Occurrent*) to a *Spatial Region*.

ρ-spatial_extent $(G, \{p_t\}) \rightarrow \{(p_t, sr)\}$

[1] we use single quotes to denote the *resource* corresponding to the person with name "John Smith"

Given:	A temporal RDF Graph G, a set of temporal paths P_t
Find:	All pairs (p_t, *sr*) such that $p_t \in P_t$ and p_t terminates with a *Spatial Entity s* and *sr* is the *Spatial Region* connected to *s* through *occurred_at* or *located_at*.
ANS ⟵ *ρ-spatial_extent* (*G*, *ρ-theme* (*G*, '101st Airborne Division'.*participates_in.Battle*))	(2)

The next two spatial operators focus on spatial relationships. As a prerequisite, we define a *spatial predicate* which is used to express conditions on spatial relationships. Spatial predicates are built from *qualitative spatial functions* and *metric spatial functions*. For a given temporal RDF graph *G*, let the set of instances of *Spatial Regions* be denoted by *S*, and let *B* denote the set {*true*, *false*}. A qualitative spatial function is a Boolean function $qsf : S \times S \rightarrow B$. Any of the following *qualitative spatial relationships* may be used as qualitative spatial functions in our formalization: *disjoint*, *meets*, *overlaps*, *covers*, *inside*, *equals*.

A metric spatial function is a function $msf : S \times S \rightarrow \mathbb{R}$. We use one metric spatial function $distance : S \times S \rightarrow \mathbb{R}$ which returns the distance between two *Spatial Regions*. We define a *metric spatial expression*, *mse*, as follows.

$$\langle mse \rangle ::= \langle msf \rangle \, \langle comp \rangle \, \mathbb{R}$$

$$\langle comp \rangle ::= \, < \mid > \mid \leq \mid \geq \mid =$$

A spatial predicate $sp : S \times S \rightarrow B$ is defined in terms of metric spatial expressions and qualitative spatial functions. It takes the following form.

$$\langle sp \rangle ::= \langle mse \rangle \mid \langle qsf \rangle \mid \langle sp \rangle \wedge \langle sp \rangle \mid \langle sp \rangle \vee \langle sp \rangle$$

The first spatial relationship operator, *ρ-spatial_locate*, is designed to retrieve entities based on their spatial relationship with a given location. An example of this type of search is "which Military Units have spatial extents which are within 20 miles of (48.45° N, 44.30° E) in the context of battle participation?" (see Example 3).

ρ-spatial_locate ({(p_t, *sr*)}, *sr′*, *sp*) ⟶ {p_t}	
Given:	A set of (*temporal path*, *Spatial Region*) pairs *X*, a *Spatial Region sr′*, and a *spatial predicate sp* defined on *sr* and *sr′*
Find:	All temporal paths p_t such that $\exists$ (p_t, *sr*) $\in X$ where *sp*(*sr*, *sr′*) evaluates to *true*
S_1 ⟵ *ρ-spatial_extent* (*G*, *ρ-theme* (*G*, *Military_Unit. participates_in.Battle*)) S_2 ⟵ (48.45° N, 44.30° E) *ANS* ⟵ *ρ-spatial_locate* (S_1, S_2, *distance* (S_1, S_2) ≤ 20 miles)	(3)

The last two spatial relationship operators, *ρ-spatial_eval* and *ρ-spatial_find* investigate how entities are related in space. These operators are based on connecting two thematic entities to their respective *Spatial Regions* and then examining the spatial relationships between those regions.

ρ-spatial_eval identifies those entities whose regions satisfy a given spatial predicate. As an example for *ρ-spatial_eval*, consider the query "which infantry unit's operational area overlaps the operational area of the 3rd Armored Division?" (Example 4).

ρ-spatial_eval ({(p_t, *sr*)}, {(p_t', *sr′*)}, *sp*) ⟶ {(p_t, p_t')}	
Given:	A set of (*temporal path*, *Spatial Region*) pairs *X*, a set of (*temporal path*, *Spatial Region*) pairs *Y*, and a *spatial predicate sp* defined on *sr* and *sr′*
Find:	All pairs of temporal paths (p_t, p_t') such that $\exists$ (p_t, *sr*) $\in X$, (p_t', *sr′*) $\in Y$ where *sp* (*sr*, *sr′*) evaluates to *true*
S_1 ⟵ *ρ-spatial_extent* (*G*, *ρ-theme* (*G*, '3rd Armored Division'. *participates_in*. *Military_Event*)) S_2 ⟵ *ρ-spatial_extent* (*G*, *ρ-theme* (*G*, *Military_Unit. participates_in.Military_Event*)) *ANS* ⟵ *ρ-spatial_eval* (S_1, S_2, *overlaps* (S_1, S_2))	(4)

ρ-spatial_find returns the qualitative spatial relationships which hold between the entities' associated *Spatial Regions*, for example "what is the spatial relationship between the 3rd Armored Division and the 101st Airborne Division in the context of military battles?" (Example 5)

ρ-spatial_find ({(p_t, *sr*)}, {(p_t', *sr′*)}) ⟶ {(p_t, p_t', *qualitative spatial relationship*)}	
Given:	A set of (*temporal path*, *Spatial Region*) pairs *X* and a set of (*temporal path*, *Spatial Region*) pairs *Y*
Find:	For each pair (p_t, *sr*) $\in X$ and (p_t', *sr′*) $\in Y$ all triples (p_t, p_t', *qsr*) for each *qualitative spatial relationship qsr* which holds between *sr* and *sr′*
S_1 ⟵ *ρ-spatial_extent* (*G*, *ρ-theme* (*G*, '3rd Armored Division'.*participates_in*. *Military_Event*)) S_2 ⟵ *ρ-spatial_extent* (*G*, *ρ-theme* (*G*, '101st Airborne Division '. *participates_in.Military_Event*)) *ANS* ⟵ *ρ-spatial_find* (S_1, S_2)	(5)

3.4 Temporal Operators

To allow queries on temporal relationships between entities and events, we first define operators for extracting the temporal properties of single entities and events and then define operators for searching relationships between the extracted temporal properties.

We define two operators for extracting the temporal properties of temporal paths: *ρ-temporal_range* and *ρ-temporal_intersect*. These two operators represent the two most basic temporal properties of a path: the interval during which an entire path is valid (*intersect*) and the interval during which a path unfolds (*range*). If we are interested, for example, in soldiers who were members of the same division during the same time period we use *ρ-temporal_intersect* (see Example 6). If we were interested in the temporal propertites of all soldiers who were members of a certain Division at any time, we use *ρ-temporal_range* (see Example 7).

ρ-temporal_intersect ({p_t}) ⟶ {(p_t, [t_1, t_2])}	
Given:	A set of temporal paths P_t
Find:	For each temporal path $p_t \in P_t$, the pair (p_t, [t_1, t_2]) such that [t_1, t_2] is the largest interval over *T* where $\forall$ (*s*, *p*, *o*)

	$\in$ *TRIPLES*(p_t) $\exists$ (s, p, o) : [t]$\in$ *TEMPORAL_TRIPLES*(p_t) for each $t\in [t_1, t_2]$, if such an interval $[t_1, t_2]$ exists

ANS $\leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (*G*, *Soldier. Assigned_to.Platoon.assigned_to. Soldier*))	(6)

ρ-temporal_range ({p_t}) $\rightarrow$ {(p_t, [t_1, t_2])}	
Given:	A set of temporal paths P_t
Find:	For each temporal path $p_t \in P_t$, the pair (p_t, [t_1, t_2]) such that $t_1 \leq t$ $\forall t \in$ *TEMPORAL_EXTENT*(p_t) and $t_2 \geq t'$ $\forall t' \in$ *TEMPORAL_EXTENT*(p_t)

ANS $\leftarrow$ *ρ-temporal_range* (*ρ-theme* (*G*, *Soldier. assigned_to*.'3rd ArmoredDivison'.*assigned_to. Soldier*))	(7)

The remaining temporal operators examine temporal relationships. To specify conditions on these relationships, we define a temporal predicate which is constructed from *qualitative* and *metric temporal functions*. For a given temporal RDF graph *G* over time domain *T*, let *I* denote the set of all time intervals over *T*, and let *B* denote the set {*true*, *false*}. A qualitative temporal function is a Boolean function $qtf : I \times I \rightarrow B$. Any of the following *qualitative temporal relationships* can be used in qualitative temporal functions in our formalization: *before*, *meets*, *overlaps*, *starts*, *during*, *finishes* and *equal*.

A metric temporal function is a function $mtf : I \times I \rightarrow \mathbb{Z}$. We use one metric temporal function $elapsed\_time : I \times I \rightarrow \mathbb{Z}$, which is defined for two disjoint time intervals as the duration of time between the end of the earliest interval and the start of the latest interval. The function returns zero if the intervals are not disjoint.

We define a *metric temporal expression*, *mte*, as follows.

$$\langle mte \rangle ::= \langle mtf \rangle \langle comp \rangle \mathbb{Z}$$

$$\langle comp \rangle ::= < | > | \leq | \geq | =$$

A *temporal predicate* $tp : I \times I \rightarrow B$ is constructed from qualitative temporal functions and metric temporal expressions. It takes the following form.

$$\langle tp \rangle ::= \langle mte \rangle \mid \langle qtf \rangle \mid \langle tp \rangle \wedge \langle tp \rangle \mid \langle tp \rangle \vee \langle tp \rangle$$

The first relationship-based temporal operator is concerned with the temporal properties of a single entity. The temporal relationship operator, *ρ-temporal_restrict,* inquires about the properties of an entity at a given time. For example, one may ask "who were members of the 101st Airborne Division during September 1944?" (Example 8) or "what battles occurred during November 1944?" The basic idea behind this operator is that we specify a thematic context and then restrict the set of result paths based on the temporal extents of those paths.

ρ-temporal_restrict ({(p_t, [t_i, t_j])}, [t_k, t_l], *tp*) $\rightarrow$ {p_t}	
Given:	A set of (*temporal path*, *time interval*) pairs *X*, a time interval [t_k, t_l], and a temporal predicate *tp* defined on [t_i, t_j] and [t_k, t_l].
Find:	Those temporal paths p_t such that $\exists$ (p_t, [t_i, t_j])$\in X$ where *tp*([t_i, t_j], [t_k, t_l]) evaluates to *true*

$I_1 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (*G*, *Soldier.assigned_to*.'101st Airborne Division')) $I_2 \leftarrow$ [09:01:1944, 09:31:1944] *ANS* $\leftarrow$ *ρ-temporal_restrict* (I_1, I_2, during (I_1, I_2))	(8)

The remaining two temporal relationship operators examine how multiple entities are related in time. These operators are designed to answer queries such as "what is the temporal relationship between supply drops and the beginning of major battles?" (Example 9) and "which speeches by President Roosevelt were given within one day of a major battle?" (Example 10).

For cases such as Example 9, we are looking for patterns between the temporal properties of one series of events or relationships and the temporal properties of a second series of events or relationships. We can think of the temporal properties of a series of events or relationships as a collection of time intervals. More specifically, for a set of (*temporal path*, *time interval*) pairs *X*, we define *TEMPORAL_EXTENT*(*X*) as the union of all intervals [t_i, t_j] such that $\exists$ (p_t, [t_i, t_j])$\in X$. Such a union of intervals is known as a union of convex intervals [13]. Ladkin identified a set of relationships between unions of convex intervals in [13]. Examples include *mostly overlaps*, *always before*, etc. The purpose of our next temporal relationship operator is to identify the union of convex intervals relationship which holds between the temporal extents of the two input sets. This operator is termed *ρ-temporal_find*.

ρ-temporal_find ({(p_t, [t_i, t_j])}, {(p_t', [t_k, t_l])}) $\rightarrow$ *union of convex intervals relationship*	
Given:	A set of (*temporal path*, *time interval*) pairs *X* and a set of (*temporal path*, *time interval*) pairs *Y*
Find:	The union of convex intervals relationship which holds between *TEMPORAL_EXTENT*(*X*) and *TEMPORAL_EXTENT*(*Y*)

$I_1 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (*G*, '3rd Armored Division'.*participates_in.SupplyDrop*)) $I_2 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (*G*, '3rd Armored Division'.*participates_in.Battle*)) *ANS* $\leftarrow$ *ρ-temporal_find* (I_1, I_2)	(9)

For cases like the example 10, we are interested in the temporal relationship between single entities and events rather than sequences of events. For analyzing these relationships, we introduce the operator *ρ-temporal_eval*. Intuitively, for two sets of temporal paths, this operator returns those pairs of paths whose temporal intervals satisfy a given temporal predicate.

ρ-temporal_eval ({(p_t, [t_i, t_j])}, {(p_t', [t_k, t_l])}, *tp*) $\rightarrow$ {(p_t, p_t')}	
Given:	A set of (*temporal path*, *time interval*) pairs *X*, a set of (*temporal path*, *time interval*) pairs *Y*, and a temporal predicate *tp* defined on [t_i, t_j] and [t_k, t_l]
Find:	All pairs (p_t, p_t') such that $\exists$ (p_t, [t_i, t_j])$\in X$, (p_t', [t_k, t_l])$\in Y$ where *tp* ([t_i, t_j], [t_k, t_l]) evaluates to true.

$I_1 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (*G*, 'Franklin Roosevelt'.*gives.Speech*)) $I_2 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme*(*G*, *Military_Unit. participates_in.Battle*))	(10)

ANS ⟵ *ρ-temporal_find* (I_1, I_2, *distance* (I_1, I_2) ≤ 1 *day*)	

3.5 Spatiotemporal Operators

Previous sections showed how we can analyze entities in the thematic and spatial dimensions and in the thematic and temporal dimensions. In this section, we show how to perform this analysis in the thematic, spatial, and temporal dimensions simultaneously. We introduce one new spatiotemporal operator: *ρ-spatiotemporal_extent* and demonstrate how to combine previously defined operators for a variety of spatiotemporal query types. We first discuss analyzing spatiotemporal aspects of single entities, which is followed by a discussion of spatiotemporal relationships between multiple entities.

The most basic spatiotemporal query on a single entity seeks to find its spatial and temporal properties for a given thematic context, for example "show the dates and locations of battles of the 101st Airborne Division" (Example 11). For such queries, we define the operator *ρ-spatiotemporal_extent*. This operator returns the *Spatial Regions* associated with (*temporal path*, *temporal interval*) pairs. It is designed to take as input the output from *ρ-temporal_range* or *ρ-temporal_intersect*.

ρ-spatiotemporal_extent (G, $\{(p_t, [t_i, t_j])\}$) ⟶ $\{(p_t, [t_i, t_j], sr)\}$	
Given:	A temporal RDF graph G, a set of (*temporal path*, *temporal interval*) pairs X
Find:	All pairs (p_t, $[t_i, t_j]$, sr) such that (p_t, $[t_i, t_j]$) ∈ X and p_t terminates with a *spatial entity* s and sr is the *Spatial Region* instance connected to s through *occurred_at* or *located_at*

I_1 ⟵ *ρ-temporal_intersect* (*ρ-theme* (G, '101st Airborne Division'*.participates_in.Battle*)) ANS ⟵ *ρ-spatiotemporal_extent* (I_1)	(11)

The *ρ-spatiotemporal_extent* operator essentially takes as input *what* and *how* (*thematic context* instance) and returns *when* and *where*. The remainder of our single-entity spatiotemporal patterns seek to answer one of *what*, *when*, or *where* given the two remaining questions and *how*, for example "when was the 3rd Armored Division within 1 mile of the coordinate (48.45° N, 44.30° E) in the context of battle participation? (Example 12)" In this example, we are searching for *when* given *what*, *where*, and *how*.

For finding when an entity is connected to a given place, we use *ρ-spatial_extent*, *ρ-spatial_locate*, *ρ-temporal_range*, and *ρ-temporal_intersect*. Intuitively, we use *ρ-spatial_extent* to obtain *Spatial Regions*. Then *ρ-spatial_locate* is used to restrict the resulting paths to those which have the desired spatial properties. Lastly, *ρ-temporal_range* or *ρ-temporal_intersect* is used to extract the temporal properties of the spatial connection.

S_1 ⟵ *ρ-spatial_extent* (G, *ρ-theme* (G, '3rd Armored Division'. *participates_in.Battle*)) S_2 ⟵ (48.45° N, 44.30° E) *ANS* ⟵ *ρ-temporal_intersect* (*ρ-spatial_locate* (S_1, S_2, *distance* (S_1, S_2) ≤ 1 mile))	(12)

For finding the location of an entity at a given time, we use *ρ-spatial_extent*, *ρ-temporal_range*, *ρ-temporal_intersect*, and *ρ-temporal_restrict*. Example 13 illustrates this for the query "where were bombing targets of the US Army Air Force in November 1943?" Intuitively, for a given set of temporal paths, we first use *ρ-temporal_range* or *ρ-temporal_intersect* to construct the temporal properties of the path. Next we use *ρ-temporal_restrict* to limit the result to the paths with the correct temporal properties, and finally we use *ρ-spatial_extent* to extract the corresponding spatial locations.

I_1 ⟵ *ρ-temporal_intersect* (*ρ-theme* (G, 'US Army Air Force'*.operates_vehicle.Vehicle.participates_in. Bombing*)) I_2 ⟵ [11:01:1943, 11:30:1943] *ANS* ⟵ *ρ-spatial_extent* (G, *ρ-temporal_restrict* (I_1, I_2, *during* (I_1, I_2)))	(13)

For finding what exists during a given time at a given Spatial *Region*, we use *ρ-spatial_extent*, *ρ-spatial_locate*, *ρ-temporal_range*, *ρ-temporal_intersect*, and *ρ-temporal_restrict*. Example 14 illustrates this pattern for the query "which soldiers were stationed within 20 miles of the coordinate (45.45° N, 37.20° E) during May 1943?" Intuitively, *ρ-temporal_range* and *ρ-temporal_intersect* first determine the temporal properties of a given path. Next, *ρ-temporal_restrict* limits these paths to those with the desired temporal properties. After that, *ρ-spatial_extent* gets the *Spatial Regions* associated with the paths, and finally *ρ-spatial_locate* restricts these paths to those with the desired spatial properties. The resulting paths represent what exists at the location and time for the given context.

I_1 ⟵ *ρ-temporal_intersect* (*ρ-theme* (G, *Soldier. stationed_at.Military Base*)) I_2 ⟵ [05:01:1943, 05:31:1943] S_1 ⟵ *ρ-spatial_extent* (G, *ρ-temporal_restrict* (I_1, I_2, *during* (I_1, I_2))) S_2 ⟵ (45.45° N, 37.20° E) *ANS* ⟵ *ρ-spatial_locate* (S_1, S_2, *distance* (S_1, S_2) ≤ 20 miles))	(14)

The remaining patterns are designed to find spatiotemporal relationships between entities. Here we are interested in when a certain spatial relationship holds between two entities, what spatial relationship holds at a given time, and what entities have a given relationship at a given time. Note that the remaining spatiotemporal patterns use generalized versions of the *ρ-temporal_range* and *ρ-temporal_intersect* operators, which take as input a set of pairs of temporal paths, and the results include the smallest time interval which contains each time point in the union of the *TEMPORAL_EXTENT* of each path in the pair or the largest time interval such that each triple in the union of the *TRIPLES* set for each path in the pair is true during the interval, respectively. The formal definitions were left out due to space constraints.

For finding when a certain spatial relationship holds we use *ρ-spatial_extent*, *ρ-spatial_eval*, *ρ-temporal_range*, and *ρ-temporal_intersect*. The basic idea is to use *ρ-spatial_extent* to find the *Spatial Region* associated with the two entities in question and to use *ρ-spatial_eval* to find the pairs of entities with the desired spatial relationship. Finally, *ρ-temporal_range* or *ρ-temporal_intersect* is used to determine the temporal properties of the two spatial connections. Example 15 illustrates this pattern for the query "when did the 101st Airborne Division come within 10 miles of the 1st Armored Division in the context of battle participation?"

$S_1 \leftarrow$ *ρ-spatial_extent* (G, *ρ-theme* (G, '101st Airborne Division'.*participates_in.Battle*)) $S_2 \leftarrow$ *ρ-spatial_extent* (G, *ρ-theme* (G, '1st Armored Division'. *participates_in.Battle*)) ANS $\leftarrow$ *ρ-temporal_intersect* (*ρ-spatial_eval* (S_1, S_2, *distance* (S_1, S_2) $\leq$ 10 miles)	(15)

For finding what spatial relationship holds at a given time, we use *ρ-spatial_find*, *ρ-spatial_extent*, *ρ-temporal_range*, *ρ-temporal_intersect*, and *ρ-temporal_restrict*. Intuitively, *ρ-temporal_range* or *ρ-temporal_intersect* is used to find the temporal properties for two entities. These entities are then filtered using *ρ-temporal_restrict*. Finally, *ρ-spatial_extent* finds the *Spatial Regions* associated with the two entities and *ρ-spatial_find* returns the spatial relationships which hold between the entities' *Spatial Regions*. Example 16 illustrates this pattern for the query "what is the spatial relationship between bombing targets of the US Army Air Force and the British Royal Air Force during November 1944?"

$I_1 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (G, 'US Army Air Force'.*operates_vehicle.Vehicle.participates_in.Bombing*)) $I_2 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (G, 'British Royal Air Force'.*operates_vehicle.Vehicle.participates_in.Bombing*)) $I_3 \leftarrow$ [11:01:1944, 11:30:1944] *ANS* $\leftarrow$ *ρ-spatial_find* (*ρ-spatial_extent* (G, *ρ-temporal_restrict* (I_1, I_3, *during*(I_1, I_3)), *ρ-spatial_extent* (G, *ρ-temporal_restrict* (I_1, I_2, *during*(I_1, I_2)))	(16)

For finding what entities have a given spatial relationship at a given time, we use *ρ-spatial_eval*, *ρ-spatial_extent*, *ρ-temporal_range*, *ρ-temporal_intersect*, and *ρ-temporal_restrict*. The basic idea is to use *ρ-temporal_range* or *ρ-temporal_intersect* to first find the temporal properties of the two entities in question. Next, *ρ-temporal_restrict* is used to filter these entities to those with the desired temporal properties. Finally, *ρ-spatial_extent* retrieves the spatial properties of the entities, and *ρ-spatial_eval* filters the entity pairs to those with the desired spatial relationship. Example 17 illustrates this pattern for the query "which US Military Units have battle areas which overlap with British Military Units' battle areas during October 1944?"

$I_1 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* (G, 'United States' .*employs_unit.Military Unit.participates_in.Battle*)) $I_3 \leftarrow$ *ρ-temporal_intersect* (*ρ-theme* ((G,'Great Britain' *employs_unit.Military Unit.participates_in.Battle*)) $I_2 \leftarrow$ [10:01:1944, 10:31:1944] $S_1 \leftarrow$ *ρ-spatial_extent* (G, *ρ-temporal_restrict* (I_1, I_3, during(I_1, I_3)) $S_2 \leftarrow$ *ρ-spatial_extent* (G, *ρ-temporal_restrict* (I_2, I_3, during(I_2, I_3)) *ANS* $\leftarrow$ *ρ-spatial_eval* (S_1, S_2, overlaps (S_1, S_2))	(17)

4. RELATED WORK

The unique aspects of our work center on the utilization of named relationships in the thematic dimension. On the web, we cannot expect complete information regarding the spatiotemporal locations of all entities. Rather, we cope with a portion of thematic entities that are associated with spatial information, specifically *Spatial Occurrents* and *Named Places*. This corresponds to our need to link non-spatial entities with spatial entities at the thematic level in order to analyze their associated spatial properties. The remainder of this sections discuss related work in ontology-based GIS and temporal GIS.

Application of ontologies in GIS focuses on practical problems of defining a common vocabulary to describe the geospatial domain which can facilitate interoperability and limit data integration problems [1, 8]. On the Web, this use of ontology for better search and integration of geospatial data and applications is embodied in the Geospatial Semantic Web [7]. We see our work as complementary to this work. The geo-ontologies created for GIS are analogous to *Named Places*, *Spatial Occurrents*, and *Spatial Regions* in our upper-level ontology. Thus, our work provides a means to further incorporate non-spatial thematic knowledge and analysis with the geospatial knowledge and analysis provided through geo-ontologies and GIS.

There is no shortage of spatiotemporal database models for temporal GIS. [15] identifies 10 distinct spatiotemporal data models. Of these, object-oriented and event-based models and the three domain model are most similar to the ontology-based approach presented here.

The three domain model, introduced by Yuan, is described in [22, 23]. This model represents semantics, space, and time separately. The semantics domain consists of concrete entities and abstract concepts which model non-spatial attributes of geographic entities. The temporal domain is composed of temporal objects representing time points and time intervals. The spatial domain is composed of simple spatial objects (e.g. points, lines, polygons) and complex spatial objects which are constructed from simple spatial objects. To represent spatiotemporal information in this model, semantic objects are linked via temporal objects to spatial objects. This provides temporal information about the semantic (thematic) properties of a given spatial region. This is analogous to temporal *occurred_at* and *located_at* relationships in our model. In the three-domain model, there is a one-to-one mapping from semantic and temporal objects to spatial objects and from spatial and temporal objects to semantic objects. This is the key difference when compared to our model because we incorporate non-spatial entities into the semantic domain and provide the notion of a thematic context to link these entities to the spatial domain in a variety of ways (many-to-many). Our Semantic Web style approach allows the incorporation and utilization of much more non-spatial information by utilizing indirect connections to the spatial domain, and it allows the direct application of existing thematic analytics techniques [4].

In [21], the authors discuss a combination of object-oriented and event-based modeling approaches for dynamic geospatial domains. They define an upper-level ontology similar to ours which distinguishes between continuants and occurrents. They also model the concept of a setting and a situate function which maps entities and events to settings. Settings can be spatial, temporal, or spatiotemporal. In contrast to our work, the authors provide a more detailed classification of event-event and event-object relationships. Our work differs in that we provide a means to assign spatial properties to those entities not directly connected to a spatial setting and in our focus on analyzing relationships. Also, no query operators are discussed in [21].

5. CONCLUSIONS

The Semantic Web community has primarily focused on the semantics of thematic data and has paid less attention to the temporal and spatial nature of data. We believe that Semantic Web technology should be enhanced by necessary models and formalizations suitable to deal with temporal and spatial data. To that end, this paper presents an upper-level ontology for modeling thematic, temporal and spatial dimensions of data, as well as formal query operators to analyze these three dimensions. The work presented here builds on our previous work on the semantic association and *ρ-queries*. The formalization of the query operations is presented in four steps. First, we introduce thematic context queries which are based on association paths. Then we present the operators for querying thematic and spatial dimensions. Next, queries on thematic and temporal aspects of data are presented. Finally, we show how queries on all the three dimensions (space, theme and time) can be expressed using combinations of operators. We show the expressiveness of the formalization by a set of example queries based on a scenario. The conceptual part of this work is presented in this paper and is a necessary precursor to our ongoing work of implementation and evaluation which we will present in the near future. We plan on creating user-defined operators in the Oracle database which build upon Oracle Spatial operators and Oracle's existing support for RDF data. Further, we will use extensible indexing [19] for efficient implementation.

6. ACKNOWLEDGEMENT

This work is partially funded by NSF-ITRIDM Award#0325464 titled 'SemDIS: Discovering Complex Relationships in the Semantic Web'.

7. REFERENCES

[1] Agarwal, P. Ontological considerations in GIScience. *Int'l Journal of Geographical Information Science*, *19*, 5 (May 2005), 501-536.

[2] Aleman-Meza, B., Nagarajan, M., Ramakrishnan, C., Ding, L., Kolari, P., Sheth, A., Arpinar, I.B., Joshi, A. and Finin, T. Semantic Analytics on Social Networks: Experiences in Addressing the Problem of Conflict of Interest Detection. In *Proceedings of the Fifteenth Int'l World Wide Web Conference (WWW '06)*, Edinburgh, Scotland, 2006.

[3] Allen, J.F. Maintaining knowledge about temporal intervals. *Communications of the ACM*, *26*, 11 (Nov. 1983), 832-834.

[4] Anyanwu, K. and Sheth, A.P. ρ-Queries: Enabling Querying for Semantic Associations on the Semantic Web. In *Proceedings of the Twelfth Int'l World Wide Web Conference (WWW '03)*, Budapest, Hungary, 2003.

[5] Bateman, J. and Farrar, S. Spatial ontology baseline. SFB/TR8 internal report I1-[OntoSpace]:D2, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany, 2004.

[6] Booch, G. *Object oriented design with applications*. Benjamin-Cummings Publishing Co., Inc., Redwood City, CA, USA, 1990.

[7] Egenhofer, M.J. Toward the Semantic Geospatial Web. In *Proceedings of the Tenth ACM Int'l Symposium on Advances in Geographic Information Systems (ACM-GIS '02)*, McLean, VA, 2002.

[8] Fonseca, F.T., Egenhofer, M.J., Agouris, P. and Camara, G. Using Ontologies for Integrated Geographic Information Systems. *Transactions in GIS*, *6*, 3 (June 2002), 231-257.

[9] Grenon, P. and Smith, B. SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition and Computation*, *4*, 1 (2004). 69-104.

[10] Gutierrez, C., Hurtado, C. and Vaisman, A. Temporal RDF. In *Proceedings of the European Conference on the Semantic Web (ESWC '05)*, Heraklion, Crete, Greece, 2005.

[11] Jones, C.B., Abdelmonty, A.I., Finch, D., Fu, G. and Vaid, S. The SPIRIT Spatial Search Engine: Architecture, Ontologies, and Spatial Indexing. In *Proceedings of the Third Int'l Conference on Geographic Information Science (GIS '04)*, Adelphi, MD, USA, 2004.

[12] Kolas, D., Hebeler, J. and Dean, M. Geospatial Semantic Web: Architecture of Ontologies. In *Proceedings of the First Int'l Conference on GeoSpatial Semantics (GEOS '05)*, Mexico City, Mexico, 2005.

[13] Ladkin, P.B. The logic of time representation, University of California at Berkley, Berkley, CA, 1987.

[14] Mukherjea, S. and Bamba, B. BioPatentMiner: An Information Retrieval System for BioMedical Patents. In *Proceedings of the Thirtieth Int'l Conference on Very Large Data Bases (VLDB '04)*, Toronto, Canada, 2004.

[15] Pelekis, N., Theodoulidis, B., Kopanakis, I. and Theodoridis, Y. Literature Review of Spatio-Temporal Database Models. *The Knowledge Engineering Review*, *19*, 3 (2004), 235-274.

[16] RDF. http://www.w3.org/RDF/.

[17] SemanticWeb. http://www.w3.org/2001/sw/

[18] Sheth, A.P., Arpinar, I.B. and Kashyap, V. Relationships at the Heart of Semantic Web: Modeling, Discovering and Exploiting Complex Semantic Relationships. in Nikravesh, M., Azvin, B., Yager, R. and Zadeh, L.A. eds. *Enhancing the Power of the Internet Studies in Fuzziness and Soft Computing*, Springer-Verlag, 2003.

[19] Srinivasan, J., Murthy, R., Sundara, S., Agarwal, N. and DeFazio, S. Extensible Indexing: A Framework for Integrating Domain-Specific Indexing Schemes into Oracle8i. In *Proceedings of the Sixteenth Int'l Conference on Data Engineering (ICDE '00)*, San Diego, CA, USA, 2000.

[20] Theodoulidis, C.I. and Loucopoulos, P. The Time Dimension in Conceptual Modelling. *Information Systems*, *16*, 3 (1991), 273-300.

[21] Worboys, M. and Hornsby, K. From Objects to Events: GEM, the Geospatial Event Model. In *Proceedings of the Third Int'l Conference on Geographic Information Science (GIS '04)*, Adelphi, MD, USA, 2004.

[22] Yuan, M. Modeling Semantical, Temporal and Spatial Information in Geographic Information Systems. in Craglia, M. and Couclelis, H. eds. *Geographic Information Research: Bridging the Atlantic*, Taylor & Francis, 1996, 334-347.

[23] Yuan, M. Wildfire Conceptual Modeling for Building GIS Space-Time Models. In *Proceedings of GIS/LIS*, Pheonix, AZ, 1994.