Analyzing Theme, Space and Time: An Ontology-based Approach[1]

Matthew Perry, Farshad Hakimpour, Amit Sheth
LSDIS Lab, Computer Science Department, the University of Georgia
Athens, GA, USA
{mperry, farshad, amit}@cs.uga.edu

Abstract

The W3C's Semantic Web Activity is illustrating the use of semantics for information integration, search and analysis. The majority of work in this community has focused on the thematic aspects of information and has paid less attention to its spatial and temporal dimensions. This poster describes ongoing work in capturing the thematic, spatial and temporal dimensions of information using an integrative ontology-based framework. We present an upper-level ontology combining the thematic and spatial dimensions and show how to incorporate temporal semantics. We also introduce the notion of a thematic context and discuss its role as a building block for querying and analysis.

1 Introduction

Relationships are what define semantics, and ontologies are the centerpiece of many current approaches to realizing semantic information systems. Additionally, semantic analytics [Aleman-Meza et al. 2006] has emerged as a powerful new class of application focusing on analysis of relationships in semantic web data (e.g., RDF graphs). Key concepts in this area are semantic associations [Anyanwu and Sheth 2003] which are complex relationships connecting resources in RDF graphs. So far, work in semantic analytics has concentrated on thematic relationships (e.g., the fact that two glycopeptides participated in the same biological process). However, in many domains and for many applications (e.g., national security, emergency response), we cannot ignore how entities are related in space and time.

To realize an extension of thematic analytics to spatial and temporal analytics, we envision a system that combines vast amounts of thematic metadata captured using semantic web data representations with digital geospatial data collected by the GIS community. In addition, we want to incorporate temporal metadata. Given the example of analyzing historical entities and events of World War II, we would seek to answer queries such as "Which military units came in close spatial proximity with the 101st Airborne Division during November 1943?" or "What water features intersect the area of Operation Market Garden?"

2 Multi-dimensional Model

We present an upper-level ontology defining a general hierarchy of thematic and spatial entity classes and associated relationships connecting these entity classes (see Figure 1). 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 when integrating thematic and spatial information, which is necessary for utilization of disparate and incomplete information sources such as data available on the web.

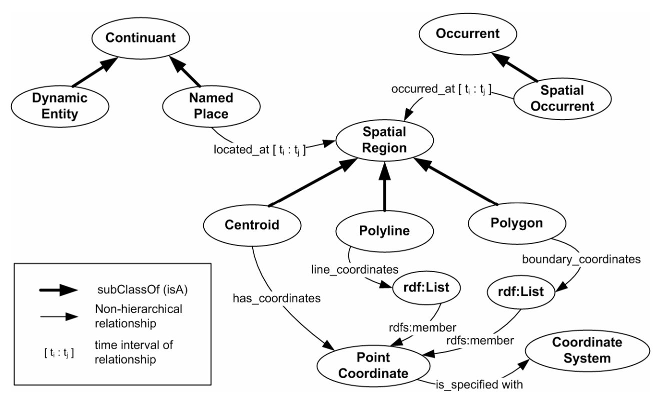

Figure 1. Upper-level ontology for space, time, and theme. *Spatial Occurrents* and *Named Places* are directly linked with *Spatial Regions*, which record their geographic location. Temporal intervals on relationships denote when the relationship holds (valid time). We intend for application-specific domain ontologies in the thematic dimension to be integrated into the upper-level ontology through subclassing of appropriate upper-level classes and relationships.

Thematic Dimension: Our upper-level thematic ontology consists of a fundamental class hierarchy and a few basic

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

relationships. We first distinguish between *Continuants* and *Occurrents* [Grenon and Smith 2004]. *Continuants* are entities that persist over time and maintain their identity through change (e.g., a soldier). *Occurrents* represent events and processes (e.g., a battle). A second division of entities concerns spatial properties. Some *Occurrents* are inherently spatial, such as a troop deployment; others are not, such as the assignment of a solider to a division. We therefore explicitly represent *Spatial Occurrents*. Similarly, 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* (e.g., a building or a city). In contrast, *Dynamic Entities* are those *Continuants* with dynamic spatial behavior (e.g., a vehicle).

Spatial Dimension: The spatial dimension of our ontology captures two-dimensional space and supports both topological relations and quantitative relations. 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*.

Inter-dimension Connections: 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 key relationships allow us to associate a thematic concept, such as the city of Berlin or the Battle of the Bulge, with its geospatial properties. Consequently, spatial relationships between thematic entities can be derived using the associated *Spatial Regions*.

Adding Temporal Support: We adopt a discrete, linearly-ordered time domain and focus on absolute time. We incorporate the time dimension by associating time intervals with relationship instances in the ontology, essentially a temporal reification of an RDF statement. The time interval on the relationship denotes the times at which the relationship is valid. This follows the approach taken in [Gutierrez et al. 2005].

3 Spatiotemporal Thematic Analytics (STTA)

The basic goal of this framework is an extension of thematic analytics which supports search and analysis of spatial and temporal relationships between entities. The key idea is that non-spatial entities 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.

Thematic Context: The most fundamental aspect in our approach to STTA is what we term a *thematic context*. The definition of a thematic context builds from the concept of ρ-path semantic associations introduced in [Anyanwu and Sheth 2003]. A ρ-path association between two resources in an RDF graph is essentially a path in the RDF graph connecting the two resources. Intuitively, a thematic context defines a *type* of association; it provides a means to constrain the thematic connections used to associate spatial and temporal properties with thematic entities. For an RDF Graph G, a thematic context is an ordered sequence $tc = C_1.P_1.C_2.P_2.C_3...C_{n-1}.P_{n-1}.C_n$ 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 ρ-path semantic association $p = e_1.p_1.e_2.p_2.e_3\ldots e_{n-1}.p_{n-1}.e_n$ 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 subclass (or descendant) of P_i 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 (or descendant) of C_i.

The connections created from thematic contexts allow a number of ways to analyze information in three dimensions. This starts with simply examining an entity's spatial and temporal properties with respect to a thematic context. Consider the example context *'3rd Armored Division'.participates_in.Battle*. By retrieving the *Spatial Regions* connected to instances of the *Spatial Occurrent* Battle, we can see the spatial properties of the 3rd Armored division with respect to a battle participation context. Similarly, examining the temporal properties of the triples making up these paths allows us to analyze temporal properties. We can also look at temporal and spatial relationships between multiple entities. With this approach, we can answer queries such as "Which military units' operational areas overlap the operational area of the 3rd Armored Division?" by retrieving the *Spatial Regions* from the previous example and finding overlapping *Spatial Regions* connected to other military units through a battle participation context. We can similarly answer temporal relationship queries such as "Which presidential speeches were given within 1 day of a major battle?" by evaluating the temporal relationships between thematic context instances.

References

[Aleman-Meza et al. 2006] Boanerges Aleman-Meza, Meenakshi Nagarajan, Cartic Ramakrishnan, Li Ding, Pranam Kolari, Amit Sheth, Ismailcem Budak Arpinar, Anupam Joshi, and Tim Finin. Semantic Analytics on Social Networks: Experiences in Addressing the Problem of Conflict of Interest Detection. *15th International World Wide Web Conference*. Edinburgh Scotland.

[Anyanwu and Sheth 2003] Kemafor Anyanwu and Amit Sheth. ρ-Queries: Enabling Querying for Semantic Associations on the Semantic Web. *12th International World Wide Web Conference*. Budapest, Hungary.

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

[Gutierrez et al. 2005] Claudio Gutierrez, Carlos Hurtado, and Alejandro Vaisman. Temporal RDF. *European Conference on the Semantic Web*. Heraklion, Crete, Greece.