Spatial Relationship Networks: Network Theory Applied to GIS Data

Barry Bitters

ABSTRACT: Throughout the short history of the digital geospatial sciences, human and natural phenomena have been modeled to develop a better understanding of complex situations. Each attempt at modeling has captured a small facet of a complex situation and created a simple representation of the interaction of forces within the complex structure. These models attempt to first identify commonly occurring associations and relationships between known features and then apply these associations to known objects and situations. If all the associations and relationships used within each of these models could be collected and captured as an assemblage of logical predicates, then, ultimately an approximation of a complex environment would result. We propose doing this as an extended form of an ontological data structure and applying the resulting network to existing cartographic feature content. The basic premise of this research is that existing geospatial databases can be augmented with additional feature content using a formal ontological database containing a extensive relationship network stored in the form of a predefined set of feature associations and their probable spatial relationships.

Ontologies are commonly used in the information technology realm to explicitly define a set of objects, concepts, or situations within a subject domain. However, ontologies usually define only what could exist, not what actually exists or what might possibly exist. Ontologies are developed with the ultimate intent of being used in some form of inference engine as a predictive tool. However, to be a truly viable predictive tool a bit more detail must be included in an ontological data structure. For a formal and explicit ontology of human and natural geographies to be a viable tool for the prediction of existence, it must provide more detail than just an itemization of a domain's class structure. Ontologies must provide some indication of the likelihood of existence – not everything that could exist; but everything that might exist and some indication of the possibility that each object would exist in a certain region or situation.

As an extension of our past work with geographical taxonomies, we are developing an extended geographic relationship network of cultural and natural features/objects. It is taking the form of a set of logical axioms – each axiom describing a specific association between two objects. In addition to the definition of an association each axiom also contains a probability factor for the occurrence of this association, a specific spatial or aspatial relationship describing this association, and a spatial probability factor.

KEYWORDS: GIS, Spatial Relationship, Association, Ontology, Semantic Web.

Introduction

Traditional methods of geospatial database production employ either the exploitation of existing sources of geospatial feature content or manual/automated digitizing of feature content from imagery. Both methods have limitations on the quantity of feature content that can be obtained. Feature content of existing geospatial databases is limited by the database specifications used to capture the data – specifications often designed for purposes other then visual simulation. In digitizing operations, economic feasibility and the question of whether the required level of detail is resolvable from the available aerial photography both dictate the feature content that will be available. Considering the current high levels of detail that are often demanded, alternative means for rapid and economical population of extremely high levels of feature detail are essential to keep abreast of ever increasing customer expectations.

The basic premise of this line of research is that there exist a wealth of untapped feature content that can be derived from existing geospatial databases – additional feature content that could be used to augment the existing feature data content used in current GIS systems. Waldo Tobler proposed the "1st Law of Geography" with the statement, "Everything is related to everything else, but nearby things are more related than distant things" [Tobler, 1970]. This bold statement has been the under-pinning premise of modern geographic information science (GISci) and most forms of spatial analysis. By applying Tobler's "1st Law of Geography" to the GIS database production process, it should be possible to significantly increase the feature content of existing feature databases.

Geographic information technology is a means for documenting the location and existence of natural and cultural features that ultimately appear in real-time simulation databases. We move about our environment and chronicle the existence of all sorts of objects. This is the gist of the data collection phase of the geographic information process. Then, using a variety of data analysis methods, we attempt to relate objects to other objects. This we call spatial modeling, and in effect is a form of after-the-fact relationship definition. This approach may be adequate for the modeling of simple systems and concepts. However, when modeling a complex adaptive system, for instance a natural or cultural landscape, we often finds this approach tends to return simplistic, and reductionist representations of more complex real-world situations.

However, if we were to approach this problem from a predictive perspective and define the relationships before the fact; it would then be possible, to a certain degree of probability, to predict the existence of new features - based purely on what is archived in geospatial databases. To do this, though, we must first create a definition of all the possible associations, spatial relationships, and their probability of existence, within a natural or complex system; this would generate a broad-based semantic knowledge database describing all those possible situations and objects that could occur within a system. This would ultimately manifest itself as a massive network of axioms that would describe the potential complex phenomena and situations within a system.

Further, because natural and cultural features and situations are regional in nature each axiom could be attributed with a variable defining a unique regional or situational setting. For instance, in the case of the street light class, there is a higher incidence of this class of feature in an urban and suburban setting than in a rural setting. In the mature urban setting, where much of the electrical infrastructure is underground, there would often be a one-to-one association of an underground access panel to each street light object. This association would normally not be present in the rural setting since rural electric infrastructure is most often overhead. In the case of the suburban setting, this association may or may not apply since the electrical infrastructure could be overhead or underground.

Associations and relationships of this sort are easily captured in the form of logical axioms with coupled levels of probability. If these axioms were to be saved as a consolidated knowledge base, the result would be an interconnected network of associations and relationships. There is a possibility when a database of this sort is applied to existing geospatial data, the location of new feature content will result – feature content that could be used to augment visual displays.

The questions we wish to explore are: Is it possible to collect commonly occurring associations and relationships and then use them to predict additional feature content for use in the preparation of real-time simulation databases? Using existing sources of vector feature data as a starting point, is it possible to logically augment that data based on probabilistic predictions? And finally, is it possible to create provisional feature databases from imagery, and then using probabilistic prediction techniques, logically augment the image-derived feature data to a more photo-realistic level of detail?

Background

Today, the creation of real-time simulation databases is nothing more than an exercise in multi-source data fusion. In its simplest form this process involves obtaining data from a variety of disparate sources, conditioning the data, registering the data to a common ground plane, harmonizing the disparate data to remove duplicate features, and ultimately post-processing the data into a form suitable for ingestion in an image generation system. The ultimate level of feature detail is therefore limited by the combined feature content of the various sources used in the data fusion process.

The current state-of-the-art of data fusion technology exists primarily as low-level object refinement, so called Level 1 data fusion. This involves combining disparate data "to obtain the most reliable and accurate estimate of an entity's position, velocity, attributes and characteristics" [Hall, 2004]. Essentially, this is the same type of processing performed in simulation database generation – the capturing of object locations and attributes. Hall further states that "Level 2 and Level 3 data fusion (situation refinement and threat refinement) are currently dominated by knowledge-based methods. …and these areas are relatively immature". Hall goes on to say, "only very primitive cognitive models exist to replicate the human performance of these functions" [Hall, 2001].

Previous research into the application of rule-based processing have shown, through the use of processes incorporating contextual content, it is possible to augment existing geospatial feature content with significant meaningful feature detail [Bitters, 2005].

The Nature of the Problem. Social network analysis is a process for the mapping and measuring of relationships and flows between people, groups, organizations, animals, computers or other information/knowledge processing entities. The nodes in the network are the people, object or groups while the links show relationships or flows between the nodes. Social network analysis provides a means for both visual and mathematical analysis of the relationships between different groups or entities. This research effort attempts to apply the basic principles of social networking to the preparation and augmentation of future GIS databases with a particular emphasis on real-time simulation databases.

The significance of this research effort is that it could:

- Develop processes and capabilities to produce probabilistic knowledge base of the feature content in natural and cultural landscapes.
- Demonstrate the power of probabilistic ontologies as a tool in modeling complex adaptive systems.
- Demonstrate the power of probabilistic ontologies to serve as a knowledge base for inference in a variety of subject domains.
- Demonstrate a common data structure to store logical descriptions of complex networks.

Semantic Web Technologies

Semantic web technology is a software capability that allows the meaning of and associations between information elements to be known and stored for subsequent use in a variety of web-based and traditional software systems [Berners-Lee, 2001]. Unlike traditional software technology, where subject matter expertise must be "hard-wired" into software applications, in semantic technology, domain knowledge, vocabularies, rules, and meanings are stored separately from data, content files, and application code [Pollock, 2004]. In semantic web technology a separate knowledge model is essential – a knowledge model describing those parts of the world to be used by software applications. The world is represented through a precise definition of how each domain object is associated with other domain objects.

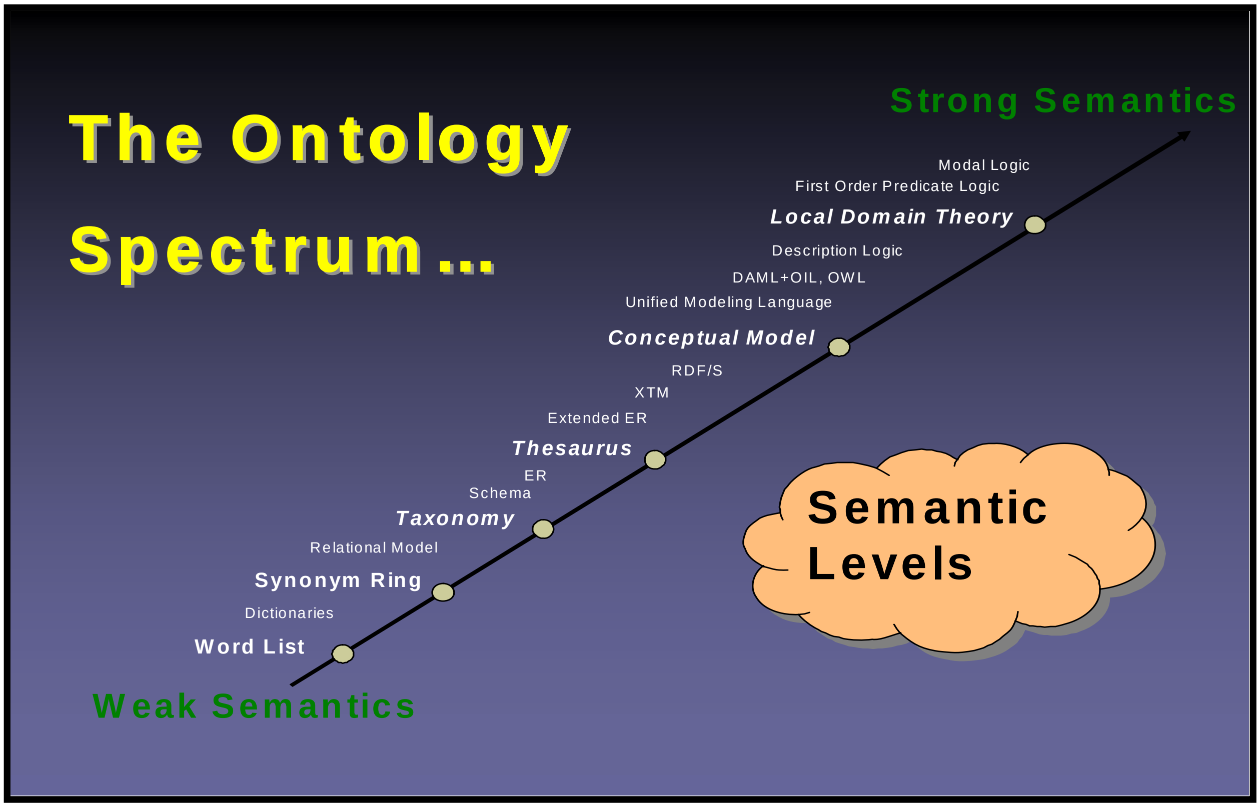

Figure 1. The Ontology Spectrum: showing the relative semantic level of various forms of ontological data structures (after Deconta 2003).

An ontology is a formal knowledge model in the form of a representation of a set of concepts within a domain of discourse and the relationships between those concepts. Ontologies can take the form of word lists, thesauri, taxonomies, and various other forms of high level conceptual models. One form of ontology is the taxonomy. It is one of the simplest forms of a knowledge model used in semantic web technologies. In taxonomies, relationships are implicitly defined by a hierarchical structure. The hierarchical structure establishes a "parent-child" relationship between domain concepts. These “parent-child” relationships are often based on functional or visual characteristics. Therefore, the structure itself becomes a means to identify the nature of basic functional relationships within a domain. Because of the hierarchical nature of all taxonomies, some concepts are often grouped under more then one category.

Spatial Associations. A spatial association is any commonly occurring co-existence between two objects. If you live in the suburbs, merely walking outside the front door will reveal some of these more subtle spatial associations. The existence of the obligatory flowering tree in the front yard is a prime example of an association - one mandated by many municipality building codes. In most modern American subdivisions, each residence will have one. A mailbox – most every residential dwelling will have one. On the edge of the right-of-way notice the utility access box – if there is public water for each residence, there will probably be a water valve in the utility box. These are all examples of commonly occurring associations – the features that spatial data mining techniques search for in the development of spatial association rules.

Spatial Relationships. A primary function of a geographic information system is determining those factors that dictate the location, distance and proximity between features; for example, the distance separating hazardous waste disposal sites from hospitals, schools, or housing developments is an example of a spatial relationship. Spatial relationships are attributes defining either absolute and/or relative locations of two or more objects. Spatial relationships can be in the form of distances and proximities between objects, direction of an object from other objects, relative movement of objects, or topological relations of two or more objects (inside, outside, intersecting, etc.). Traditional geographic information science (GISci) is concerned with a limited set of spatial relationship such as Equal, Disjoint, Intersects, Touch, Overlap, Cross, Within, and Contains. However, as can be seen in Table 1, there are significantly more spatial relationships in common use than the few traditionally used in GISci.

Table 1. Commonly Used Spatial Relationships

Spatial Term	*Spatial relationship*
above	Write your name above the line.
across	The house is across the street.
against	She leans against the tree.
ahead of	The truck is ahead of the car.
along	The river bank is along the river.
among	He is standing among the trees.
around	The fence is around the yard.
behind	The shed is behind the house.
below	Write your name below the line.
beneath	He sat beneath the tree.
beside	The girl is standing beside the boy.
between	She is between two trees.
from	He came from the house.
in front of	The mail box is in front of the house.
inside	He is inside the house.
nearby	There is a tree nearby the house.
off	His hat is off.
out of	He came out of the house.
through	She went through the door.
toward	She is walking toward the house.
under	He is hiding under the table.
within	Please mark only within the circle.

Topological Relationships. Topology is the mathematics of connectivity and adjacency. Topological relationships are the most common relationship found in spatial databases. In GIS, topological relationships are used to describe spatial relationships between points, lines and areas – point connectivity between line segments at nodes (connectivity) and what exists on each side of line segments (adjacency). Topology is a powerful tool for the determination of relative spatial positions. However, topological relationships are not the only relations that exist to describe common relation among geographic features.

Functional Relationships. Functional relationships are spatial properties describing commonly occurring associations that exist between features – connections and interactions that exist in the real world. Functional relationships not only include the obvious; for instance, the existence of a highway intersection at the location where two highways cross; but also include the more subtle situations like the existence of paired traffic control devices (stoplights, stop signs, or yield signs) located at most highway intersect.

Logical Relationships. Functional relationships are usually defined using a logical statement in the form of a logical relationship. Logical relationships describe dependencies – conditions that may exist that can dictate the existence of another object, a situation, or an event - based on the existence of some known object, situation, or event. Logical relationships are often described using the “if-then”, “and-or”, or “not” conditional statements.

Plausibility and the Relationship Situation. Current GISci is based on the assumption that we can inventory geographic features and then from this inventory predict the existence of other objects, situations and events. However, the real-world is a network of complex interacting factors. It is an over-simplification to think that the complex nature of any social, economic, or natural environment can be adequately defined using only a narrow set of logical and functional relationships. The real-world can not be adequately described merely by using a set of "if-then" relationships. Because some level of uncertainty is always present, when modeling the real-world, this uncertainty factor must be incorporated as an integral factor of the "if-then" computational process.

Ontologies as Semantic Knowledge Bases. In information technology "Ontologies communicate a common understanding of a domain, declare explicit semantics, make expressive statements, and support sharing of information" [Lacy, 2005]. In the future semantic web, ontologies will provide authoritative descriptions of a domain – descriptions of both objects and the way objects interact with each other. This will serve as the knowledge base for all advanced processing on the World Wide Web.

However, the standard ontological form used to describe natural and cultural objects is inadequate, especially when ontologies are employed in any predictive function. Merely defining a set of classes, with definitions, and hierarchical relationships provides only a superficial representation of knowledge within a domain – a representation that is adequate only in very basic data mining operations. However, if ontologies are to be an effective definition of domain knowledge for complex systems and networks, then the complex nature of the system must be incorporated into the ontology - not merely as a definition of nodes and links within a complex network, but as an effective predictive tool – a tool that represents all the uncertainties inherent in a particular complex environment.

Most forms of data analysis are based on a certain level of uncertainty. Therefore, it is essential that some measure of probability be built into classes and attributes of classes within the typical ontology. In a geographical ontology, a more appropriate way to describe natural and cultural features would be to expand the traditional ontological form to include weighted probabilities for all relationships, associations, and characteristics. In this way, the ontology would not only define absolute existence; it would define the probability that an object might exist within a defined environment. In this way, it would then be possible to later infer the plausibility of existence. For this to be possible though, qualified causal relations, dependencies, and interdependencies must be included as an integral element of all data components of an ontology. These must be in the form of explicit values and states of probability.

OWL/RDF. The Resource Description Framework (RDF) is a family of specifications and a general method for modeling information [Manola, 2004]. RDF offers a standardized semantic network data model that can be further formalized by ontology modeling languages such as RDF Schema (RDFS) and the Web Ontology Language (OWL). RDF and its related technologies are currently used extensively as a means to document domains of computer science. OWL is a family of languages for storing ontologies. Designed for information processing, OWL was developed to be read and understood by semantic web applications. With a larger vocabulary and stronger syntax than that of RDF, OWL allows interpretation and processing by a wide variety of computer applications. Both RDF and OWL are expressed using Extensible Markup Language (XML).

Relationship Networks. In social network analysis, associations and relationships are viewed in terms of *nodes* and *links.* Nodes are the individual actors/objects within the networks, and links are the associations and relationships between the actors/objects. Social networks have been used to characterize complex sets of associations and relationships between members of various types of social systems [Brandes, 2005].

An extensive geographical ontological data structure exists that identifies 18,000 classes of natural and cultural objects that commonly appear in the real world [Bitters, 2007]. Expanding this ontology to include detailed spatial relations, associations, and probabilities would create an advanced logical network of knowledge – a relationship network - for use in automated image analysis, intelligence analysis, and for geospatial production in general.

Stored as an ontological data structure in OWL/RDF format, this relationship network would provide a means to describe and store many of the complexities of both natural and cultural geographies. Using this network of knowledge as a means for the representation of uncertainty, within a logical inference engine (classical, Bayesian, Dempster-Shafer, or some future form of logical inference engine) it would be possible to determine probable existence of additional detail – in other words - to extract information and conclusions that would otherwise be missed using more conventional logic tools.

Methodology

The following processes are necessary to create a usable association's ontology of natural and cultural objects:

- Identify object namespace.
- Define spatial relationship classes.
- Design ontological data structure
- Develop association database
- Develop/acquire rudimentary inference engine.
- Test inference engine.

Initially, the process of creating an ontology of associations and their spatial relationships is an effort in observation – observing features in the real-world and recording those commonly-occurring associations that exist. It also involves consulting references to identify associations that have been documented in past research. The process also includes the capture of local, national, and international laws, ordinances, rules, and common practices that affect the placement of features on the Earth's surface. The information obtained from each of these research efforts must finally be reviewed by subject matter experts to insure the validity of the observed associations, spatial relations and probabilities.

Object Namespace. In a formal ontological data structure, associations and spatial relations are defined using a controlled vocabulary to insure a consistent and unambiguous semantic is employed. For a project of this nature to be a feasible undertaking, a comprehensive object namespace must be available to allow the use of standard feature class naming. The Visual Objects Taxonomy (VOT) [Bitters, 2007] was chosen because it contains a hierarchical data structure of over 18,000 named and defined natural and cultural features. It was developed as a standardized data structure for use in the creation of visual databases for use in visual simulation. The VOT data structure is an ideal candidate for this project because it is currently being translated into a set of formal OWL/RDF ontologies and will soon be available for general public use.

Spatial Relationship Class Definition. Prior to the design of a formal ontological data structure it was essential to define a set of spatial relationship classes – a set of relationship classes that could be used to describe a wide selection of common place spatial relations. It would serve as a means to store a comprehensive set of functional relations describing probable interactions of natural and cultural features to each other. These spatial relations, in conjunction with their respective associations would then serve as a mechanism to attach quantifiable values to each association statement.

Table 2. Defined spatial relationship classes.

Basic Relationship	**Containment Relationship**
srLocationOf( )	srOutOf( )
srFrom( )	srContains( )
srToward( )	srWithin( )
Over/Under Relationship	srPartOf( )
srAbove( )	srSurround( )
srBelow( )	srPartialSurround( )
srOnTop( )	srAmong( )
srBeneath( )	srConsistsOf( )
srOver( )	srInside( )
srUnder( )	srOutside( )
srOff( )	**Orientation Relationship**
srOn( )	srOrientation( )
Adjacency Relationship	srParallel( )
srAdjacent( )	srParallelOutside( )
srAcrossFrom( )	srParallelWithin( )
srBetween( )	srParallelPartialWithin( )
srAlong( )	srPerpendicular( )
srAlongside( )	srPerpendicularOutside( )
srSide( )	srPerpendicularWithin( )
srRightSide( )	srFacing( )
srLeftSide( )	**Intersection Relationship**
srAround( )	srIntersects( )
srAttached( )	srTouches( )
srBackSide( )	srOverlaps( )
srBehind( )	srThrough( )
srFront( )	srCrosses( )
srAhead( )	**Surface Relationship**
srInFront( )	srOnGround( )
srAgainst( )	srAboveGround( )
srFacing( )	srUnderGround( )
Proximity Relationship	**Network Relationship**
srNear( )	srBranchOf( )
srFar( )	srTributaryOf( )
Containment Relationship	srTraverses( )
srCongruent( )	srInterconnecting( )
srConnectedTo( )	**Cluster Relationship**
srConnectWith( )	srCluster( )
srDetached( )	srClusterRandom( )
srHasPart( )	srClusterLinear( )
srOutOf( )	srClusterNonRandom( )

Ontology Design. The explicit definition of a feature association and their spatial relations must at a minimum include the following critical elements of information:

- A reference object
- An associated object
- An association probability
- A spatial relationship
- A relationship probability
- A regional delimiter
- A temporal delimiter (optional)
- An explicit logic statement (optional)

The Reference Object is the object for which an association or relationship is being defined. It is identified using an explicit reference to a unique and defined feature class – a class previously defined in an existing ontology.

The Associated Object is the feature class that is associated with the discrete reference object in an association statement. It is identified using an explicit reference to a unique and defined feature class – a class defined in an existing ontology. An example of an association is a house and a land parcel. Houses are usually built on a parcel of land. (Exception: in some areas of the world houses are built on stilts over water and have no associated land parcel.)

The Association Probability is a value between 0.01 and 1.0 identifying an objective (sometimes subjective) evaluation of the likelihood of an association to occur in the real world. Table 3 provides a list of values and approximate verbal equivalents. As an example, a house is most often associated with a land parcel. Therefore, a probability of 0.99 would be assigned. (This value could vary based on the region of the world.)

The Spatial Relationship defines the spatial characteristic of an association. Table 2 provides an extensive set of classes describing spatial relationship for inanimate, non-sentient features. Each relationship class can include a variable that further defines the spatial characteristic. As an example the proper form of the proximity relationship – *srNear* is *srNear(10.0)* meaning the reference object and the association object are within 10.0 meters of each other. Multiple relationships may be defined for each association where each spatial relationship describes a different spatial aspect of the association. In a relational database, each relationship would be stored as a separate record. As an example, a house is usually located on a land parcel - *srOn(0.0).*

The Relationship Probability is a value between 0.01 and 1.0 identifying an objective (sometimes subjective) evaluation of the likelihood of a spatial relationship to be valid in the real world. Table 3 provides a list of values and approximate verbal equivalents. As an example, a house is most often on a land parcel (*srOn* from Table 3). Therefore, a probability of 0.99 would be assigned. This value could vary based on the region of the world.

Table 3. Probability/Frequency Qualifiers

Probability	*Percentage*	*Descriptor*
0.99	100	Always
0.9	90	Usually
0.8	80	Regularly
0.7	70	Often
0.6	60	Frequently
0.5	50	Sometimes
0.4	40	Occasionally
0.3	30	Infrequently
0.2	20	Seldom
0.1	10	Rarely
0.01	00	Never

The Regional Delimiter describes a region or country of the world for which the association and relationship applies. The visual appearance of a natural or cultural landscape is a region-specific phenomenon. From an overhead, aerial view, landscapes in different regions often look alike. However, on the ground, those same locations can often look dramatically different and associations and relationships will vary accord to local laws, traditions, customs, construction practices, and a variety of other factors.

An optional Temporal Delimiter is available to record any temporal characteristics that might affect an association or spatial relationship. As an example, this field might be used to indicate the temporal nature of on-street parking of automobiles in residential areas.

An optional Explicit Logic Statement can be used to further expand on the spatial relationship. For example in the case of the mailbox/street association, the *srNear(1.0)* function could be expanded to *(Width + srNear(1.0))* to formally state that the mailbox location would be the street width plus 1.0 meters from the street centerline.

Accuracy of Probability Values. Classical and more rigorous forms of inference are dependent on the *a priori* assignment of representative probability factors. For an association axiom to be usable in an inference engine, it must have a probability factor assigned that rates the likelihood of occurrence in any particular situation. Rarely will these values be objective representations of the actual occurrence of any particular association or relation. It must be assumed that initial attempts at defining these probability values will include personal and systematic biases. Tuning these values to remove potential biases will likely be more time consuming then documenting the initial spatial relationships as axiomatic expressions.

Spatial Association Definition. Laws, ordinances, common practices, habits, and human nature dictate most spatial associations on the cultural landscape. In natural environments, spatial associations tend to appear to be more random. However, this appearance of randomness must not be misconstrued as a lack of predictable associations. In natural settings, though less distinct, associations do exist and can be qualified; if not quantified.

Each association could have multiple relationships to describe all possible spatial aspects of the association as is illustrated in Table 4. Using the single-family residential dwelling (SFRD) and mailbox as an example the following associations and spatial relationships would describe the typical situation of a mailbox associated with a detached single-family residence in a suburban setting:

- A mailbox would be positioned between the SFRD and the street.
- The Mailbox would face the street.
- The Mailbox would be near the street - within 1.0 meters.
- If a driveway were present, the mailbox would be within 3.0 meters of the driveway.
- If a walkway from the house to road is present the mailbox would be within 1.0 meters of the walkway.

Table 4 shows a relational database representation of these associations and relationships and their probabilities. Figure 2 illustrates the same values in OWL/RDF format. Notice that each association can be expressed as a statement of fact. For example, in a suburban setting a mailbox can be associated with a SFRD. Further, ~90% of the time the mail box will be present. Three spatial relationships are defined for the SRFD-mailbox association:

1) The mailbox is between the SRFD and the street.
2) The mailbox is facing the street.
3) The mailbox is near the street (within 1.0 meter).

Secondary and tertiary associations can provide insight into additional spatial relationships. The location of a driveway or walkway relative to a SFRD will provide added information to further refine positional information for the mailbox. A driveway or walkway often connects a house to a street. In a suburban setting, residential mailboxes are often positioned a short distance from these paved surfaces. By augmenting the primary SFRD/mailbox association with the spatial relation information associated with secondary and tertiary associations it is often possible to refine the position of features. In the SFRD/mailbox example in Table 4, if a walkway is present, a mailbox will be within 1.0 meters of the walkway. If a driveway is present, a mailbox will be within 3 meters of the driveway. This is an example of networked relationships and how they can sometimes be used to refine inferences and derived spatial locations.

Table 4. Associations and spatial relationships describing the single-family residential dwelling (SFRD) and the Residential Mailbox in a suburban setting.

Reference Object	*Association Object*	*Association Probability*	*Spatial Relationships*	*Probability*
SFRD	**Mailbox**	**0.9**	**srBetween(Street)**	**0.99**
			srFacing(Street)	**0.99**
SFRD	**Street**	**0.8**	**srFacing(Street)**	**0.99**
			srInFront(Street)	**0.99**
SFRD	**Street**	**0.9**	**srConnect(Driveway)**	**0.99**
Mailbox	**Street**	**0.99**	**srNear(1.0)**	**0.99**
Mailbox	**Driveway**	**0.8**	**srNear(3.0)**	**0.99**
Mailbox	**Walkway**	**0.5**	**srNear(1.0)**	**0.99**

Often spatial associations and their spatial relationships appear to be intuitively obvious. At the same time, they are enlightening because they establish a discrete geometric correlation between the location and orientation of associated objects relative to the reference objects. In the example above, a spatial connectivity is established between the street and the SFRD via a driveway. These are all significant photo-identifiable signatures and in the future have the potential to be autonomously identified from imagery.

Results

Figure 3 shows a graphic representation of the preliminary result of this experiment. The intent was to determine if, using a limited set of feature associations, it would be possible to predict the location of new feature content and augment existing geospatial data. Notice that the predicted locations of new feature content are consistent with possible locations of the representative types of features that were tested. However, the predicted feature positions are only approximate and do not represent the actual and exact ground positions of their real-world counterparts. Figure 4 illustrates the method used to compute the location of each new feature. Associations relating to the single-family residence, the driveway and the street feature were found in the relationship network. The spatial extent of each association was determined as follows: The blue area represents the “BETWEEN” spatial relation – the area between the house and the road. The green area represents the “NEAR” spatial relationship – the area near the driveway and in front of the house. The red area represents the “NEAR” spatial relationship – the area near the road and in front of the house. The geometric intersection of all of these selected spatial relationships is shown in are black and is the area in which the mailbox would probably be located. Based on this technique, a new mailbox feature could then be added to a database at the centroid of the geometric intersection.

A limited set of associations and relationships have been created in OWL/RDF format. These were used in a rudimentary inference engine to evaluate the efficacy of the probabilistic feature augmentation concept. This relationship ontology was applied to a set of local government GIS data for a small residential subdivision. Street centerlines and building footprints were used in this test case. Cumulative probability computations were not performed since only a limited set of associations had been entered into the association ontology. The inference algorithm did however consider a minimum allowable probability value for selection. New feature content was generated during this effort. However, the following discrepancies were encountered in the feature predictions:

Figure 3. An aerial photo of a suburban setting overlaid with existing geospatial data (building footprints and street centerlines in black) and predicted positions of probabilistic features (residential mailboxes, and stop signs in red and manholes in yellow.) Image courtesy of http://seamless.usgs.gov.

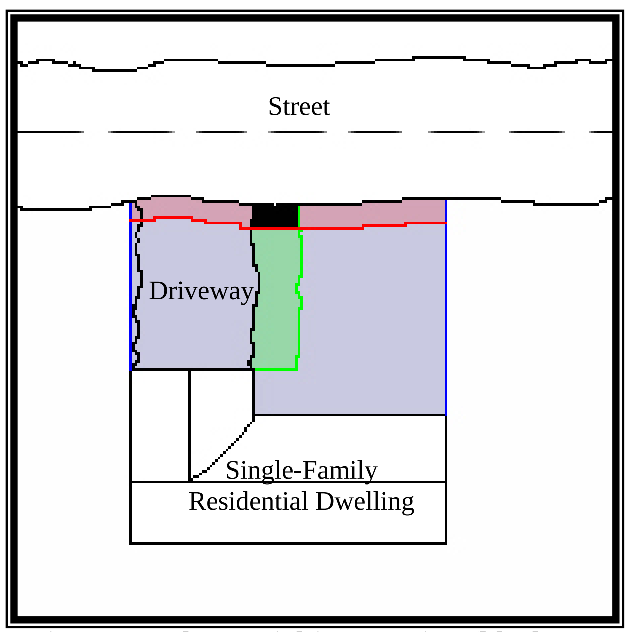

Figure 4. The spatial intersection (black area) of the spatial relationships of all feature associations - used to predict the location of new feature content.

```
<owl:Class rdf:ID=”SFRD”>
    <sAssoc:type>Mailbox
        <assocProb>0.9</assocProb>
        <sRelate:type>srBetween
            <sRAttrib:type>Street</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
        <sRelate:type>srFacing
            <sRAttrib:type>Street</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
    </sAssoc:type>
</owl:Class>

<owl:Class rdf:ID=”SFRD”>
    <sAssoc:type>Street
        <assocProb>0.8</assocProb>
        <sRelate:type>srFacing
            <sRAttrib:type>Street</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
        <sRelate:type>srInFront
            <sRAttrib:type>Street</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
    </sAssoc:type>
</owl:Class>

<owl:Class rdf:ID=”SFRD”>
    <sAssoc:type>Street
        <assocProb>0.9</assocProb>
        <sRelate:type>srConnect
            <sRAttrib:type>Driveway</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
    </sAssoc:type>
</owl:Class>

<owl:Class rdf:ID=”Mailbox”>
    <sAssoc:type>Street
        <assocProb>0.99</assocProb>
        <sRelate:type>srNear
            <sRAttrib:type>1.0</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
    </sAssoc:type>
    <sAssoc:type>Driveway
        <assocProb>0.8</assocProb>
        <sRelate:type>srNear
            <sRAttrib:type>3.0</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
    </sAssoc:type>
    <sAssoc:type>Walkway
        <assocProb>0.5</assocProb>
        <sRelate:type>srNear
            <sRAttrib:type>1.0</sRAttrib:type>
            <sRProb:type>0.99</sRProb:type>
        </sRelate:type>
    </sAssoc:type>
</owl:Class>
```

Figure 2. The single-family residential dwelling mailbox example encoded into OWL/RDF format.

- More manholes were predicted than actually exist on the ground. A more detailed algorithm is required for the prediction of locations for this feature class.
- Fewer stop signs were predicted than exist on the ground. Algorithmically determining priority at intersections is problematic.
- In this locale, not all houses have residential mailboxes. Determining which houses do not have mailboxes is algorithmically difficult.

Conclusions and Future Research

This preliminary research effort indicates that there is potential for the determination of additional feature content from existing geospatial databases. With additional study, this approach to spatial database production has the potential to allow the addition of new feature content to geospatial databases. This will only be possible after the development of a broad-based ontology of natural and cultural features and their commonly occurring associations and their spatial relationships. To date, a broad-based OWL/RDF ontology defining extensive associations and relationships is not yet available. Early prototyping has shown that the inference of new feature locations from *a priori* spatial relationship information is possible. It is possible to predict reasonably accurate ground locations of new feature content. However, the precise positioning of new feature content based on stored spatial relationships is still problematic and warrants further study.

Many potential uses exist for spatial relationship networks. The development of spatial databases to support future very high detail visualization environment would be an ideal candidate for this type of database production. The deployment of military systems in tactical environments is based on repetitive exercise – repetitive exercise based on soldiers “training the way they fight”. The rich friendly and opposing force doctrine provides many of the associational and inter-relation information necessary to generate a spatial relationships network – a database that could have broad implications in future battle space prediction. This type of database could also be used in the data fusion process to fill gaps in intelligence collected and generated from non-persistent surveillance systems. It could also be used as an automated mentoring tool to train image analysts. The analyst would identify an object on an image and the software could suggest other features that might be associated with the selected feature. In the environmental sector, the spatial relationship network could also be used as a tool to model the inherent intricacy of complex adaptive natural systems.

GIS is concerned with the creation and storage of detailed data describing real-world environments – environments that are complex adaptive systems. In the past, GIS systems were limited by computing capabilities and as a result produced and stored simplistic representations of environments. In the future, with anticipated improvements in computing and display capabilities, it will be possible to portray the true complexities of cultural and natural environments. Only by harnessing the complex network of relatedness within real world environments, it will be possible to portray and model these environments in their actual depth of detail.

REFERENCES

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The Semantic Web: A new form of Web content. ***Scientific American*.** May, pp 35-43.

Bitters, B. (2005). Enhanced urban synthetic environments through logical feature population techniques. ***Proceedings of the Image 2005 Conference,*** Scottsdale, AZ, July 2005.

Bitters, B. (2007). ***Universal Feature Classification Taxonomy and 3-D Model Library*** http://vissim.uwf.edu/ . Last visited 01 June 2008.

Brandes, U. & Thomas E. (Eds.). (2005). ***Network Analysis: Methodological Foundations***. Heidelberg: Springer-Verlag.

Costa, P. C. G. (2005). ***Bayesian Semantics for the Semantic Web*.** Doctoral Dissertation. Department of Systems Engineering and Operations Research, George Mason University: Fairfax, VA, USA. p. 312.

Daconta, M. C., Obrst, L. J. & Smith, K. T. (2003). ***The Semantic Web: A guide to the Future of XML, Web Services and Knowledge Management.*** Indianapolis, IN: Wiley Publishing, Inc.

Hall, D. L., & Llinas, J. (2001). ***Handbook of Multisensor Data Fusion,*** Boca Raton, Florida: CRC Press LLC.

Kessler, O. (1992) ***Functional Description of the Data Fusion Process.*** Warrington, PA: Office of Naval Research, Naval Air Development Center.

Lacy, L. W. (2005). ***OWL: Representing Information Using Web Ontology Language***, Victoria, BC, Canada: Trafford Publishing.

Manola, F. and Miller, E. (2004). ***RDF Primer: W3C recommendation*** http://www.w3.org/TR/2004/REC-rdf-primer-20040210/ . Last visited 01 June 2008.

Pollock, J. T. & Hodgson R. (2004). ***Adaptive Information: Improving Business Through Semantic Interoperability, Grid Computing, and Enterprise Integration***. New York: J. Wiley and Sons,.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region**.** ***Economic Geography*** 46 (2):234-40.

Barry Bitters, Senior Scientist, Science Applications International Corporation, 8301 Greensboro Rd P.O. Box 50132, McLean, Virginia 22102, bittersb@saic.com