A Linguistically Grounded Annotation Language for Spatial Information

James Pustejovsky – Jessica Moszkowicz – Marc Verhagen

Department of Computer Science
Brandeis University
415 South Street
Waltham, MA
USA

{jamesp,jlittman,marc}@cs.brandeis.edu

ABSTRACT. *An understanding of spatial information in natural language is necessary for many computational linguistics and artificial intelligence applications. In this paper, we discuss what problems researchers face with respect to this topic, focusing on the need for a well-developed annotation scheme. The desiderata for such a specification language are defined along with what representational mechanisms are required for the specification to be successful. We then review several spatial information annotation schemes, focusing on the latest version of the ISO-Space specification. Finally, we discuss where ISO-Space still falls short and propose some ways that the specification could be enriched.*

RÉSUMÉ. *Algorithmes dans les deux la linguistique computationnelle et en intelligence artificielle nécessitent la capacité de comprendre et de raisonner sur l'information spatiale en langage naturel. Dans cet article, nous discutons de ce que les chercheurs face à des problémes en ce qui concerne ce sujet, mettant l'accent sur la nécessité d'un schéma d'annotation bien développé. Les desiderata d'un tel langage de spécification sont définies avec les mécanismes de représentation sont nécessaires pour la spécification de réussir. Nous examinons ensuite plusieurs schémas d'annotation d'information spatiale, en se concentrant sur la derniére version de la spécification ISO-Space. Enfin, nous nous demandons oú ISO-Space est encore loin et de proposer quelques façons que les spécifications peuvent être enrichis.*

KEYWORDS: *Spatial relation annotation, ISO-Space, motion representation*

MOTS-CLÉS : *Annotation relation spatiale, ISO-Space, représentation du mouvement*

1. Introduction to the Problem

The automatic recognition of spatial information in natural language is currently an area of considerable attention in computational linguistics and Artificial Intelligence. The development of algorithms that exhibit "spatial awareness" promises to add needed functionality to NLP systems, from named entity recognition to question answering and text-based inference. In order for such systems to reason spatially, however, they require the enrichment of textual data with the annotation of spatial information in language. This involves a large range of linguistic constructions, including spatially anchoring events, descriptions of objects in motion, viewer-relative descriptions of scenes, absolute spatial descriptions of locations, and many other constructions.

In this paper, we describe the motivations for and the specification of an annotation language for encoding spatial and spatiotemporal information as expressed in natural language texts. We describe the syntax and application of ISO-Space, a language designed for this domain. Throughout the exposition of the representations, we illustrate what inferences are enabled by the language, as well as what capabilities are presently outside the scope of the specification.

To illustrate the application of the proposed annotation language, let us consider several use cases in greater detail. The first example highlights the demands on a spatial annotation language, as it involves the identification and interpretation of places, geolocations, and objects in motion, as reported in a travel blog, such as the one below [1].

> John left San Cristobal de Las Casas four days ago. He arrived in Ocosingo that day. The next day, John biked to Agua Azul and played in the waterfalls there for 4 hours. He spent the next day at the ruins of Palenque and drove to the border with Guatemala the following day.

In addition to the motion information supplied by predicates such as *leave* and *bike* and the explicit references to locations (e.g., *San Cristobal de Las Casas*), an interpretation of this excerpt requires a fairly robust understanding of the temporal information at play as well. It has already been demonstrated that the annotation of events, times, and their relative ordering from natural language text is greatly beneficial in aiding subsequent inference and reasoning over natural language texts (Setzer, 2001; Mani and Wilson, 2000; Pustejovsky *et al.*, 2003; Denis and Muller, 2011). TimeML (Pustejovsky *et al.*, 2005) was designed with just such applications in mind. Extending this paradigm to space, SpatialML (Mani *et al.*, 2008; Mani *et al.*, 2010) has provided a robust platform for the subtask of geolocating geographic entities and facilities in text. The logical next step is to resolve toponyms in the text, when there is uncertainty or ambiguity, a problem addressed in Wing and Baldridge (2011). An additional challenge presented by texts such as this, however, is that the motion, the

1. This text was taken from http://www.rideforclimate.com, but was slightly edited for explanatory purposes.

moving objects, and the paths along which these motions occur, must be identified in order to understand the spatial consequences expressed by the text. In this particular case, the problem involves tracking John's movement as he biked through Central America.

A different challenge presents itself with the second use case, which involves a dynamically evolving situation that is being witnessed by multiple viewers, each reporting with multiple feeds, such as Twitter or SMS texts, with both geolocated and time-stamped messaging capabilities. Such data may be created during an event such as the Occupy Wall Street movement, but one can easily imagine a use case on a smaller scale in which a few friends are trying to meet up at a particular location. For example, consider the following SMS exchange:

A: Just arrived. Where are you?
B: Arriving at T stop from Park Station.
A: I'm at the east entrance.
B: How do I get there?
A: Head left out of the train and I'll meet you.

The issues of spatial representation and reasoning from language in such a use case are quite complex. While the above example is relatively simple, in general, such a use case could involve multiple relative frames of reference, integration of time-indexed reports, the normalization or registration of spatial data associated with geolocated objects mentioned in the text or referenced by accompanying images. In direction giving contexts, deictic updating of the local interpretation of landmarks and relative frames of reference is an even more difficult problem (Prévot, 2004). Muller and Prévot (2009) point out the role that different feedback strategies play in how spatial values are grounded in the interpretation of the discourse [2].

A related use case involves the integration of geosensor data over a monitored area with verbal utterances from viewer feeds. A spatial annotation language, to be useful for such cases, must minimally capture each of these dimensions of spatial information, and allow for this data to supplement or complement information coming from other modalities representing contextual or spatial information. The viewer descriptions in this example are similar in many respects to the street-level annotations provided in the PURSUIT corpus (Blaylock, 2011).

A simplified version of this situation is schematically represented in Figure 1, where two viewers are supplying feeds about the activities and position of an individual in motion, each relative to their point of view. The observers are associated with time-stamped observations in natural language. The locations of the observers are known, either relatively or absolutely in terms of their GPS coordinates. In the latter case a map can overlay the entire scene. The boxes represent the outlines of two images taken relatively shortly after each other, from different angles. The person is

2. Cf. recent work by Levit and Roy (2007) and Vogel and Jurafsky (2010).

represented twice in the scene, once at time t1 and once at time t2. The only other object in the scene is a newsstand [3].

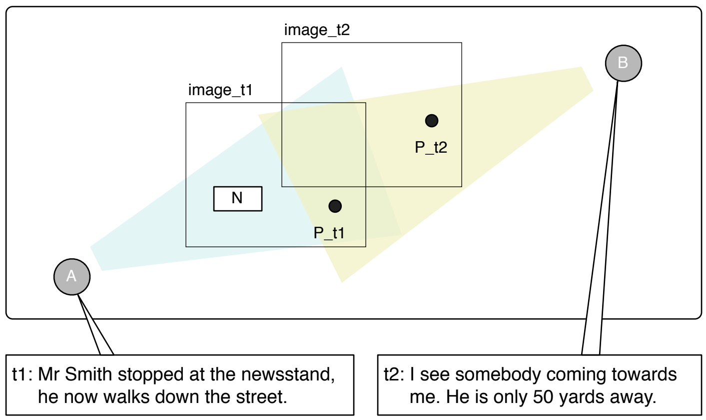

Figure 1. *Multiple-viewer object tracking*

A third use case involves the interpretation of directions for moving toward a specific goal, as illustrated below:

> Walk towards the main entrance of the building and turn left to exit. You will see steep stairs directly in front of you. Go up these stairs and follow the path.

Cognitive psychologists have long been interested in the language of route descriptions, in part because they presuppose such rich conceptual spatial schemata (Taylor and Tversky, 1992; Klippel *et al.*, 2005). The issues raised by this example primarily involve the interpretation of orientation, the identification of landmarks as used for navigation, as well as the fact that the viewpoint is being updated as the directions are traversed by the reader (Denis, 1997). In fact, it is often the case that the supplier of such directions is depending on the reader being at a specific location in order to interpret a subsequent direction. For instance, in the example, the author does not explicitly say where the path will be in relation to the stairs. Perhaps the path actually begins when the mover reaches the middle of the staircase, but this is so clear in context that the author does not feel it necessary to supply this detail. [4]

The final use case involves integrating verbal spatial descriptions with metric data derived from identification of landmarks through geoname referencing. The goal is to use "viewer perspective" linguistic descriptions of an image to help geo-locate the viewer, when no geolocating software or capability is available to the viewer/user.

3. There is considerable work on integrating depictions of spatial configurations with descriptions (Tversky and Lee, 1999), a somewhat related problem to the present one.
4. In GIS systems, wayfinding routes can often be expressed as sequences of landmarks and the paths connecting them (Raubal and Winter, 2002; Duckham M., 2010). The importance of being able to integrate landmark metric information with descriptions emerges below as well.

This involves identifying landmark spatial configurations relative to the viewer. Landmarks are visually identifiable objects with fixed spatial locations, and are typically large man-made or physical structures (e.g. buildings, communication antennas, hills) that play an important role in navigation and wayfinding decisions. Figure 2 is an example of such an image with three landmarks identified: Horticultural Hall, the Prudential Center, and the Christian Science Office Building in downtown Boston [5].

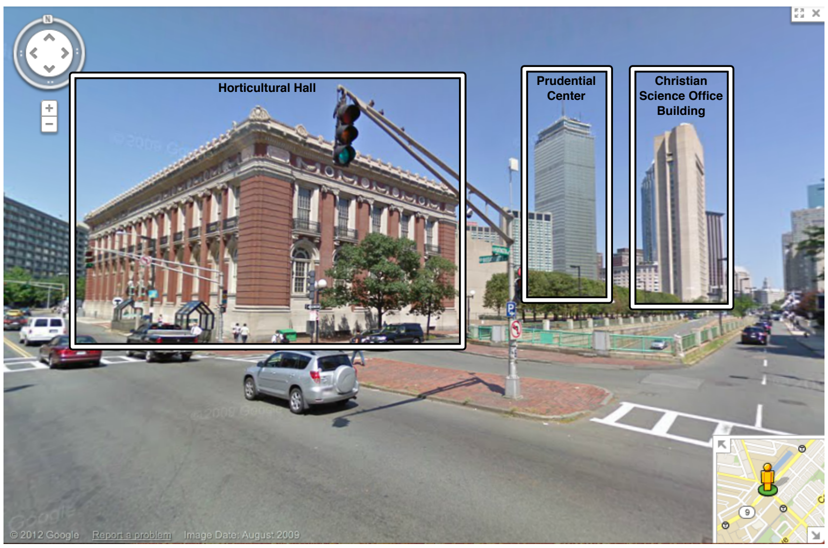

Figure 2. *Image labeling*

By interpreting the spatial descriptions of the viewer relative to such landmarks, we would be able to help in identifying the viewer's orientation and general location. The spatial annotation language should be rich enough to encode the relative and intrinsic reference frames, while also capturing and exploiting positional values that may be available from GPS designations, either from the image capture metadata, or geolocation information from gazetteers.

Given the above use cases, we now describe the semantic requirements for the annotation language, and then discuss the structure of the annotation framework. We then outline the basic elements of the current version of ISO-Space, followed by explicit examples of how these elements are used for markup. We briefly discuss our strategy of corpus-driven development of the specification, and conclude with remaining issues and outstanding questions of interpretation.

5. Image annotation has been the focus of attention within the context of content-based image retrieval (CBIR). In particular, the work done within the TRIPOD project at Sheffield has examined the different ways that geo-referenced images can be described (Edwardes *et al.*, 2007).

2. Desiderata for the Specification Language

Given the discussion and the use cases outlined above, it is clear that the specification language for spatial information in language will need to support the following computational tasks:

(1) a. identification of the appropriate topological configuration between two regions or objects; e.g., containment, identity, disjointness, connectedness, overlap, and closure over these relations, when possible;

b. identification of directional and orientational relations between objects and regions, including the distinction between frames of reference;

c. identification of metric properties of objects and metric values between regions and objects, when possible; e.g., distance, height, and width;

d. identification of the motion of objects through time and a characterization of the nature of this movement;

e. provide clear interoperable interfaces to existing representations and geodatabases, e.g., Geonames, ArcGIS, Google Earth.

Requirement (1ea) is central to all spatial information tasks. Historically, topological relationships have been the focus of much of the existing work in the field of qualitative reasoning, resulting in the development of several qualitative spatial calculi, such as the Region Connection Calculi (Randell *et al.*, 1992), the work of Asher and Vieu (1995), and the Intersection Calculi (Egenhofer and Franzosa, 1991; Kurata and Egenhofer, 2007). The specification language must incorporate such qualitative spatial relationships between regions or objects while also recognizing the need for additional non-topological relationships.

The second requirement in (1e) details what non-topological relationships must be included. To account for the data presented in the use cases from Section 1, the specification must account for many different kinds of spatial relationships in addition to the standard topological ones including directional and orientational relations, such as those discussed in Freksa (1992), Ligozat (1998), Mitra (2004), and Renz and Mitra (2004).

Desideratum (1ec) is closely related to the second requirement in that it refers to additional properties that spatial objects can have in relation to one other. Given the complexity of the use cases described earlier, the specification language must capture as much metric detail as possible between regions and objects in space, as reviewed in Cohn and Renz (2001).

In addition, the specification language must account for the motion of objects over time in a direct way. We saw that, in many of the use cases, motion is an important aspect of spatial information. In some ways, motion is easier to identify than the complex relationships that spatial objects may share, but motion can also have a lasting effect on the interpretation of a text with respect to spatial information. For that reason, the specification must include a detailed characterization of the nature of a movement.

The final desideratum calls for the specification language to be interoperable with existing resources. This is a pragmatic requirement for the specification that may, at times, be at odds with the other desiderata, which mostly involve the expressiveness of the specification. While the specification should be robust enough to account for the information described by the first four requirements, existing representations and geodatabases such as Geonames and Google Earth already do a good job of representing some aspects of spatial information, particularly geolocations. The specification language must not preclude connections to these resources and, in fact, it should be designed to take advantage of them whenever possible.

To satisfy the desiderata above, any specification language will require at least the following representational mechanisms:

(2) a. the representation of locations as regions in an interpreted spatial domain;
b. the representation of objects as occupying regions of space;
c. the identification of appropriate topological values within the model;
d. the representation of direction and orientation, both for the domain of discourse and for regions and objects;
e. the representation of metric properties;
f. the representation of object movement.

3. Current Approaches and Models

The specification language discussed in Section 4 is motivated by the desiderata and representational mechanisms introduced in the previous section. Before describing the most recent version of ISO-Space, however, we want to discuss several earlier worthwhile attempts at capturing spatial information. In this section, we discuss three of these specifications: SpatialML, the initial version of ISO-Space, and Spatial Role Labeling.

3.1. *SpatialML*

In SpatialML, locations are annotated with PLACE tags and mapped to data from gazetteers. Topological and orientation relations among locations are also represented to some extent. The annotation scheme is intended to be language-independent, and annotated corpora in English as well as Chinese have been released [6], along with editing tools [7] and automatic tagging software [8]. SpatialML focuses mainly on geographic locations, and can easily be integrated with GML, the Geography Markup Language defined by the Open Geospatial Consortium. It is also automatically mapped to

6. http://www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2011T02,
http://www.ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2010T09.
7. http://callisto.mitre.org, http://sourceforge.net/projects/spatialml/files/latest/download.
8. http://sourceforge.net/projects/spatialml/files/MIPLACE-release-v1.0b.tar.gz/download .

Google Earth's Keyhole Markup Language (KML), allowing visualization of geographic locations and relations in Google Earth.

SpatialML marks named and nominal mentions of locations, with the former being provided with geo-coordinates where possible. Locations are classified into one of twenty different types. Example (3) shows a named location with its geo-coordinates and the ISO country code for Mexico, along with a nominal location that corresponds to a body of water. Named locations are disambiguated using a gazetteer. Locations are also classified into one of 20 type codes.

(3) The next day, John biked to Agua Azul and played in the waterfalls there for 4 hours.
The next day, John biked to <PLACE id='1' form='NAM' country='MX' type='PPL' latLong='17.2625, -92.1202' >Agua Azul</PLACE> and played in the <PLACE id='2' form='NOM' type='WATER' >waterfalls</PLACE> <SIGNAL id='3' >there</SIGNAL> for 4 hours.
<LINK id='4' source='1' target='2' signals='3' linkType='IN' />

In addition to representing named places as points, SpatialML also represents topological relations between places treated as regions. The above example also shows that the waterfall and Agua Azul are topologically related by IN, meaning a tangential or non-tangential proper part (TPP and NTPP respectively in RCC8). The reason for the collapsing of TPP and NTPP is that annotators find it difficult to determine whether the part touches or does not touch the border of the including region. The RCC8 inverse links TPPi and NTPPi are not included in SpatialML, since they can be annotated by swapping arguments. The rest of the RCC8 relations are in SpatialML, whose complete set of topological relations is shown in Table 1.

Table 1. *Topological relations in SpatialML*

relation	**example**
IN (tangential and non-tangential proper parts)	[Paris], [Texas]
DC (disconnection)	the [well] outside the [house]
EC (external connection)	the border between [Lebanon] and [Israel]
EQ (equality)	[Rochester] and [382044N 0874941W]
PO (partial overlap)	[Russia] and [Asia]

As shown in example (4), SpatialML also represents a limited number of orientation relations, captured in the values of the PLACE tag's mod attribute (23 codes) and the RLINK tag's direction attribute (20 codes). These values include cardinal directions and some orientation relations (ABOVE, BEHIND, etc). Frames of reference are differentiated in SpatialML as EXTRINSIC, INTRINSIC, and VIEWER, corresponding to the absolute, intrinsic, and relative frames of reference respectively.

(4) A building five miles east of Fengshan.
A <PLACE id='1' form='NOM' type='FAC' >building</PLACE>
<SIGNAL id='2' type='DISTANCE'>five miles</SIGNAL>
<SIGNAL id='3' type='DIRECTION'>east</SIGNAL> of
<PLACE id='4' form='NAM' country='TW' type='PPL'
latLong='22.6333,120.35'>Fengshan</PLACE>.
<RLINK id='5' source='4' destination='1' distance='2' direction='E'
frame='VIEWER' signals='2 3' />

Overall, SpatialML provides substantial coverage of place names and their mapping to geo-coordinates, as well as nominal references to places. It provides some support for qualitative reasoning in 2D about topological relations, and it also captures, somewhat spottily, a small number of orientation relations. Note that the problem of qualitative reasoning involving mixed granularities, e.g., map points as in SpatialML (or polygons) and qualitative regions as in RCC8, remains to be explored.

3.2. *ISO-Space 1.0*

While SpatialML does well with 2D topological relationships and, to some extent, orientational relationships, its primary focus on geolocations neglects to account for the locations of events and motion in general. The initial ISO-Space specification attempted to build on SpatialML while incorporating spatiotemporal information by way of TimeML for event information and Spatiotemporal Markup Language (STML) for capturing motion concepts.

ISO-Space 1.0 (Moszkowicz and Pustejovsky, 2010) had at its core a LOCATION tag that is used for any area of space that is important to the annotation, meaning it will participate in some kind of spatial relationship. LOCATIONs can be static places such as *Boston* or coerced locations such as *car*. In this initial conception of ISO-Space, moving objects including individuals could also be annotated with the LOCATION tag. The specification also allowed for implicit locations to be introduced with this tag. For example, the region that is *behind the store* could be captured as a LOCATION. The attributes for the LOCATION tag were largely inherited from SpatialML's PLACE tag, though the exact set of attributes was never finalized.

The inclusion of motion in ISO-Space 1.0 was a driving force behind much of the specification. The L_PATH tag (location path) was introduced to capture geographic paths such as *the road*. The E_PATH tag (event path) was used whenever a motion event was identified in the text. Event paths were defined by a start LOCATION and an end LOCATION, at least one of which had to be specified. For example, given the text *John left Boston*, the motion predicate *left* would introduce an E_PATH with an underspecified end location and a start location of Boston. Both L_PATH and E_PATH participate in a link tag called MOTION_PATH, which connects the motion predicate to a particular path.

Spatial prepositions were also important additions to ISO-Space 1.0. A spatial signal tag was used to capture words that relate an event to a particular location or a motion to a particular path. Some prepositional phrases were considered spatial functions, however, if they introduced an implicit location or path as in *There was an accident in front of the bank.* In this case, the expression *in front of* was said to introduce a new LOCATION that was then used as the location for the accident event. S_FUNCTION tags always introduced a new LOCATION in ISO-Space 1.0, even if the output location was explicitly named in the text (e.g., *the store next to the bank*). A qualitative spatial link (QSLINK) was then used to detail how the newly introduced location fit into the annotation.

The need for spatial functions was largely motivated by the way in which ISO-Space 1.0 linked elements of the annotation to one another. The original specification called for three kinds of links: MOTION_PATH, EVENT_LOCATION, and QSLINK. The first of these was exclusively used to relate a motion event to either an event path or a location path. The second link type could only be used to connect a non-motion event to a location. Finally, qualitative spatial relationships were captured with QSLINK, which could only take locations as arguments. QSLINK was used in ISO-Space 1.0 for both topological and relative relationships between locations. Topological QSLINKs allowed for the same relations as SpatialML's LINK tag. A relative QSLINK, on the other hand, was given a relation type that was derived from the spatial function prepositional phrase that prompted the link.

Section 4 presents a substantially updated version of ISO-Space in which many of the problems of this initial specification are explicitly addressed. First though, we examine one of ISO-Space's contemporaries: Spatial Role Labeling.

3.3. *Spatial Role Labeling*

Kordjamshidi and her colleagues approach the task of spatial information representation as one of *spatial role labeling* (Kordjamshidi *et al.*, 2010). The annotation scheme they employ involves identifying and classifying certain elements of a text as spatial arguments and then relating these arguments to each other. Spatial role labeling is analogous to semantic role labeling. The related specification is referred to as Holistic Spatial Semantics (HSS) because the complete utterance rather than an isolated word is the main unit of analysis. In practice, this means that annotation is performed at the sentence level.

The HSS annotation scheme uses four tags that capture spatial arguments and an additional tag that pulls these arguments together. The TRAJECTOR and LANDMARK tags are respectively similar to the figure and ground elements in a spatial relation. One of the attributes of LANDMARK, however, introduces a major difference between HSS and ISO-Space. This attribute, called `path`, is primarily relevant when the sentence involves motion. The inclusion of motion events is an important addition to any spatial information annotation scheme, though ISO-Space and HSS differ greatly on how this information is represented.

In HSS, every LANDMARK is related to a path. If this path is deemed irrelevant, the path attribute is given a value of ZERO. This is generally the case when motion is not involved. When motion is present, one of three possible `path` values is used: BEGIN, MIDDLE, and END. Example (5) shows a case of an END path.

(5) John went into the room.
<LANDMARK id='1' path='END' >the room</LANDMARK>

Significantly, HSS puts path information in the LANDMARK tag, which effectively labels the LANDMARK as the source, goal, or midpoint of a motion. This in stark contrast to the ISO-Space approach, which aims to reflect a compositional semantics in which the path an individual traverses during a motion is elevated to a first class status in the annotation.

Motion events and spatial relation words are also explicitly captured in HSS using the MOTION-INDICATOR and SPATIAL-INDICATOR tags, respectively. The latter of these tags is where the spatial relationship between the TRAJECTOR and LANDMARK is defined. Table 2 displays the various attributes and possible values for SPATIAL-INDICATOR. Note that the annotator first chooses a value for `general-type`. The value for this attribute limits the possible values for `specific-type`, which, in turn, limits the values for the `spatial-value` attribute.

Table 2. *Attributes for SPATIAL-INDICATOR*

general-type	**specific-type**	**spatial-value**
REGION	RCC8	EC, DC, EQ, PO, TPP, NTPP, TPPi, NTPPi
DIRECTION	ABSOLUTE	S, W, N, E, NE, NW, SE, SW
	RELATIVE	LEFT, RIGHT, FRONT, BEHIND, ABOVE, BELOW
DISTANCE	QUALITATIVE	Predefined terms like *near, far*
	QUANTITATIVE	Numbers and values in text form for the key distance information

When the `general-type` is region, the annotator chooses between the RCC8 relations. For example, the SPATIAL-INDICATOR *in* in *The girl is in the room* is annotated with the `spatial-value` TPP. Such an annotation does not seem to capture the desired interpretations of the sentence. It is not really clear what is meant by the topological relationship TPP, which is generally used to relate 2-dimensional objects. Given that, the annotation employed in HSS may even suggest that the girl is somehow clinging to the side of the room rather than the presumedly intended interpretation in which the girl's feet are touching the floor of the room. The latest version of ISO-Space, described in Section 4, aims to ameliorate these problems by enriching the ways in which spatial relations are annotated [9].

SPATIAL-INDICATOR carries the bulk of the work in HSS as far as relating the TRAJECTOR and LANDMARK arguments goes. While an additional tag called SR actually pulls together the tagged elements in a sentence, SPATIAL-INDICATOR is the only place where the nature of the spatial relationship can be described. This restriction and the inclusion of the `path` attribute in LANDMARK forces the HSS scheme to annotate motion events with both the MOTION-INDICATOR and SPATIAL-INDICATOR tags as shown in example (6).

(6) She went to school.
<TRAJECTOR id='1'>She</TRAJECTOR>
<LANDMARK id='2' path='END'>school</LANDMARK>
<SPATIAL-INDICATOR id='3' general-type='REGION'
specific-type='RCC8' spatial-value='TPP'>to</SPATIAL-INDICATOR>
<MOTION-INDICATOR id='4'>went to</MOTION-INDICATOR>
<SR id='5' trajector='1' landmark='2' spatial-indicator='3'
frame-of-reference='INTRINSIC' motion-indicator='4'/>

Notice that *to* is included in the extent of the motion tag as a SPATIAL-INDICATOR with a somewhat curious `spatial-value` of TPP. Given that the LANDMARK is labeled as the end of the path, this annotation essentially puts the TRAJECTOR in the LANDMARK much in the same way that earlier examples captured *The girl is in the room*. This again, though, seems to overlook much of the complexity of spatial language that ISO-Space strives to capture.

4. Enriching the Language: ISO-Space 1.4

ISO-Space 1.0 was developed following a meeting at Brandeis University in 2009 and then refined at two workshops in 2010 and 2011. It is now in its fourth incarnation and substantial changes have been made in an effort to meet the representational requirements described earlier in this paper [10]. As before, ISO-Space 1.4 incorporates the annotations of static spatial information, borrowing from the SpatialML scheme, and events, borrowing from the TimeML scheme.

9. Note that despite SpatialML's reliance on RCC8 for the majority of its relations, that specification is not subject to this problem because only geolocated places are tagged. Therefore, SpatialML would not mandate any relationship between the entities in the example sentences (i.e., *the girl* is not a SpatialML PLACE). By allowing for more complex spatial objects than 2D regions, HSS and ISO-Space must also account for the complex relationships such objects may share.

10. The core working group includes, besides the authors: Harry Bunt, Kiyong Lee, Inderjeet Mani, and Annie Zaenen. For a description of version 1.3, see Pustejovsky *et al.* (2011).

ISO-Space is being developed as a Work Item within the ISO/TC37/SC4. It assumes the ISO CD 24612 Language Resource Management - Linguistic Annotation Framework standard (Ide and Romary, 2004). As such, ISO-Space provides a stand-off annotation scheme with some tags in the specification linked explicitly to text offsets and others representing relationships between other tags. Further, it assumes the distinction between an abstract and concrete syntax for the definition of the annotation specification, as argued in Bunt (2010). This allows for a clear semantics defined for the abstract syntax, independent of implementation.

In this section, we review the ISO-Space 1.4 elements and discuss in what ways this version is an improvement over version 1.0. We begin with the basic tags of ISO-Space: the tags that include text offsets (in most cases) and describe the basic spatial elements within a document. We then turn to the link tags, which capture more complex spatial information by relating the basic elements together.

4.1. *Basic ISO-Space Elements*

4.1.1. *Location Tags*

Locations in ISO-Space come in two varieties: PLACE and PATH. Each of these tags captures a specific kind of spatial information in the text and they can both subsequently participate in spatial relationships by way of the link tags defined below.

PLACE Tag

The PLACE tag is inherited almost directly from SpatialML. This tag is used to annotate geographic entities like lakes and rivers, as well as administrative entities like towns and counties. A PLACE tag in ISO-Space must generally be directly linked to an explicit span of text, though the specification does not preclude the use of this tag to represent implicit locations. Some examples of this tag are presented in (7). Note that additional spatially relevant elements in these sentences are left unmarked for now; only the PLACEs are shown.

(7) a. A fishing trawler swept away more than a year ago by a tsunami off [**the east coast of Japan**$_{pl1}$] has been spotted floating near [**British Columbia**$_{pl2}$], Canadian officials said Friday.

The attributes for the PLACE tag are largely inherited from SpatialML.[11] For example, for those places that have known latitude and longitude values, the `latLong` attribute can be used to allow for mapping to other resources such as Google Maps. ISO-Space also includes the *Document Creation Location* or `DCL` attribute. This is a special place that serves as the “narrative location”. If the document includes a DCL, it is generally specified at the beginning of the text, similarly to the manner in which

11. In fact, given a SpatialML annotation, an ISO-Space annotation should simply be able to inherit the captured PLACE elements.

a Document Creation Time (DCT) is specified in TimeML. If a place is the DCL, this is marked with a special attribute in the annotation of the place. The current set of PLACE attributes is shown in Figure 3.

`id`	pl1, pl2, pl3, ...
`type`	BODYOFWATER, CELESTIAL, CIVIL, CONTINENT, COUNTRY, GRID, LATLONG, MTN, MTS, POSTALCODE, POSTBOX, PPL, PPLA, PPLC, RGN, ROAD, STATE, UTM
`form`	NAM or NOM
`continent`	AF, AN, AI, AU, EU, GO, LA, NA, PA, SA
`country`	a two letter ISO 3166 country code See `http://www.iso.org/iso/country_codes/iso_3166_code_lists/`.
`state`	a principal subdivision of a country like state, province or parish, again following ISO 3661.
`county`	a subdivision below the state level
`ctv`	CITY, TOWN or VILLAGE
`gazref`	gazetteer name plus a colon plus an identifier, e.g. IGDB:2104656
`latLong`	a coordinate from the gazetteer
`mod`	a spatially relevant modifier
`dcl`	true or false

Figure 3. *Attributes for PLACE tag*

The values for the `type` attribute are currently identical to the values from the SpatialML PLACE tag with the exception of some types such as VEHICLE, which is a spatial named entity in ISO-Space 1.4, and ROAD, which is a path. Future versions of ISO-Space, however, will likely make more use of spatial ontologies such as GUM as a way to classify places (cf. Bateman *et al.*[2010]). Places can be in the form of proper names (*New York*) or nominals (*town*), which are marked with the `form` attribute as NAM or NOM, respectively. Examples of what constitutes each `type` can be found in the complete ISO-Space Specification, available at `www.iso-space.org`.

The `mod` attribute for PLACE is present to capture cases such as *tall building*, *the higher observation deck* and *two towers*, where *tall*, *higher* and *two* do not really constrain the location of the entity but they do add spatial information. In ISO-Space, this attribute is substantially different from its counterpart in SpatialML where it was used for modifiers like *bottom of the well*, *Burmese border*, *near Harvard*, *northern India* and *the right side of the building*. In many cases, these modifiers were deemed necessary in SpatialML because it focuses on annotating gazetteer entries. In ISO-Space, these cases are analyzed in two ways: (i) the SpatialML modifier is a signal for spatial relations or (ii) the entire phrase is a place [12].

12. Given this discrepancy with SpatialML, it is likely that the ISO-Space annotator will have to perform some "clean-up" of the PLACE elements that are inherited from a SpatialML annotation. This issue would be more a matter for the annotation guidelines, though, and is not relevant to the present discussion.

`id`	p1, p2, p3, ...
`beginID`	identifier of a location tag
`endID`	identifier of a location tag
`midIDs`	list of midpoint locations, if specified
`form`	NAM or NOM
`gazref`	gazetteer name plus a colon plus an identifier, e.g. IGDB:2104656
`latLong`	a coordinate from the gazetteer
`mod`	a spatially relevant modifier

Figure 4. *Attributes for PATH tag*

The ISO-Space 1.4 PLACE tag captures much of what the LOCATION tag in version 1.0 of the specification was designed to capture. The current version of the specification, however, is actually closer to SpatialML's PLACE tag in many ways. The ISO-Space PLACE tag does not capture moving spatial objects or entities that are coerced to locations in the way that the old LOCATION tag did. Such elements are represented in version 1.4 using other tags. Another change is that the PLACE tag is really designed to capture explicit spans of text rather than implicit locations. There is nothing in the specification, though, that prevents the annotator from creating a PLACE that has a null offset. In fact, allowing for implicit PLACE tags to be introduced may be necessary for certain inference tasks, so it is expected that future versions of ISO-Space will include so-called "non-consuming" PLACE tags.

PATH Tag

A PATH is a location where the focus is on the potential for traversal or its function as a boundary. This includes common nouns like *road* and *river* and proper names like *Route 66* and *Kangamangus Highway*. Paths typically have begin points and end points, although these are often not expressed in the text. Example (8) shows an instance of a PATH in which the endpoints happen to be explicit [13].

(8) the [**rail road**$_{p1}$] from [**Boston**$_{pl1}$] to [**Maine**$_{pl2}$]
Path(p1, beginPoint=pl1, endPoint=pl2)

The attributes of PATH are a subset of the attributes of the PLACE element, but with the `beginID` and `endID` elements added as shown in Figure 4. As in version 1.0 of the specification, paths are first class objects in ISO-Space. However, the distinction between an event path and a location path is no longer part of the specification. The

13. Formal treatments within qualitative spatial reasoning typically encode paths as lines, with end points (or boundaries) designating begin and end locations, e.g., Galton (1995). Without temporal indexing and ordering, of course, there is no distinction possible between the two points, so most of these theories adopt a first-order model of temporal interpretation.

new PATH tag is similar to the L_PATH tag from version 1.0 in that it is used to capture explicit locations in which traversal is the focus. What was once stored in E_PATH is represented within the new MOVELINK, described below.

4.1.2. *Non-Location Tags*

Location tags essentially designate a region of space that can be related to other regions of space. In addition, ISO-Space 1.4 allows for non-location elements of a text to be coerced into behaving like a region of space so that they may participate in the same kinds of relationships. There are three of these kinds of tag in ISO-Space: SPATIAL_NE, EVENT, and MOTION. Note that, for the most part, annotating these tags should not be the responsibility of the ISO-Space annotator. Instead, capturing this kind of information should be left to other annotation schemes and it will be left to the ISO-Space annotator to recognize when such an element should participate in an ISO-Space link tag or if additional information specific to spatial language needs to be added to the annotation. Details on this will be available in the ISO-Space annotation guidelines.

SPATIAL_NE Tag

A SPATIAL_NE is a named entity that is identified as participating in an ISO-Space link tag. The example in (9) shows which named entities in the text are considered SPATIAL_NE tags.

(9) The new [**tropical depression**$_{sne3}$] was about 430 miles (690 kilometers) west of the southernmost Cape Verde Island, forecasters said.

When a named entity is identified as a SPATIAL_NE, it receives an `id` attribute for the ISO-Space annotation [14] and the annotator may add additional attributes as suggested in Figure 5 [15]. Note that in the original ISO-Space, spatial named entities were captured with the overloaded LOCATION tag. This made it exceedingly difficult to determine what attributes the LOCATION tag should incorporate. ISO-Space 1.4 alleviates much of the burden on the annotator by making spatial named entities a separate element in the annotation. It can still be confusing to distinguish between a PLACE and a SPATIAL_NE in ISO-Space 1.4; this is an issue that must be addressed in the annotation guidelines or in subsequent versions of the specification.

Non-Motion EVENT Tag

An EVENT is a TimeML event that does not involve a change of location but is directly related to another ISO-Space element by way of a link tag. Events are inherited directly from a TimeML annotation and require no further specification in ISO-Space.

14. In lieu of being assigned an ISO-Space `id` number, the subsequent link tag can use the ID that was previously assigned to the entity by the layered annotation scheme.
15. The decision to do this is likely task-based. The annotation guidelines for a specific task will instruct the annotator on what attributes to add, if any.

`id`	sne1, sne2, sne3, ...
`form`	NAM or NOM
`latLong`	a coordinate
`mod`	a spatially relevant modifier

Figure 5. *Attributes for SPATIAL_NE tag*

`id`	m1, m2, m3, ...
`motion_type`	MANNER or PATH
`motion_class`	MOVE, MOVE_EXTERNAL, MOVE_INTERNAL, LEAVE, REACH, DETACH, HIT, FOLLOW, DEVIATE, STAY

Figure 6. *Attributes for MOTION tag*

Whereas ISO-Space 1.0 required the annotator to capture spatially relevant events explicitly with an ISO-Space tag, ISO-Space 1.4 strives to take advantage wherever possible from a layered annotation approach. That is, since TimeML already does a good job of capturing events and almost all events are located in space, there is no need to annotate them again in ISO-Space. Instead, a TimeML event can participate directly in an ISO-Space spatial relationship without being tagged in the ISO-Space annotation explicitly.

MOTION Tag

A MOTION is a TimeML event that involves a change of location. Since motions are inherently spatial, they play a special role in ISO-Space. When a TimeML event has been identified as a MOTION, it gets re-annotated with the attributes given in Figure 6.

The `motion_type` attribute refers to the two distinct strategies for expressing concepts of motion in language: *path constructions* and *manner-of-motion constructions* (Talmy, 1985). This is illustrated in the sentences in (10), where m indicates a manner verb, and p indicates a path. In the first sentence, the motion verb specifies a path whereas in the second the motion verb specifies the manner of motion. Both are annotated as motions since the motion is implied in the manner-of-motion verb.

(10) a. John arrived$_p$ [by foot]$_m$.
b. John hopped$_m$ [out of the room]$_p$.

Motion classes are taken from Pustejovsky and Moszkowicz (2008), which was based on the motion classes in Muller (1998). These classes are associated with a spa-

`id`	s1, s2, s3, ...
`cluster`	identifies the sense of the preposition
`semantic_type`	DIRECTIONAL, TOPOLOGICAL

Figure 7. *Attributes for SPATIAL_SIGNAL tag*

tial event structure that specifies, amongst other things, the spatial relations between the arguments of the motion verb at different phases of the event.

Identifying a TimeML event as a MOTION is the first step in the semantically grounded ISO-Space treatment of motion. This tag along with the MOVELINK tag are primarily motivated by the treatment of motion described by the Dynamic Interval Temporal Logic (Pustejovsky and Moszkowicz, 2011), which accounts for motion predicates at the sentence level, and Dynamic Location Update theory (Moszkowicz, 2011), which examines the impact that motion predicates can have on the interpretation of subsequent sentences in a discourse.

4.1.3. *SPATIAL_SIGNAL Tag*

A SPATIAL_SIGNAL is a relation word that supplies information to an ISO-Space link tag. It is typically a preposition or another function word or phrase that reveals the particular relationship between two ISO-Space elements.

Spatial signals are treated in a substantially different way in ISO-Space 1.4 than in version 1.0. Rather than splitting spatial prepositions among two tags, version 1.4 opts for only one tag that captures the way in which two ISO-Space elements are related. Spatial functions are no longer given a different treatment than simple spatial signals in this version of the specification. So as not to lose valuable information about the preposition, however, the set of attributes for SPATIAL_SIGNAL has been expanded, as shown in Figure 7.

Sense information, which is stored in the `cluster` attribute is optional. The values for this attribute come from a sense inventory of spatial prepositions that is currently being constructed. The `semantic_type` attribute helps the annotator decide, along with sense information if it is available, what kind of ISO-Space relationships the signal triggers. Some examples of typical SPATIAL_SIGNALs are shown in (11).

(11) a. The book is [**on**$_{s1}$] the table.
spatial_signal(s1, cluster="on-1", semantic_type=topological, directional)

b. Boston is [**north of**$_{s2}$] New York City.
spatial_signal(s2, cluster="north_of-1", semantic_type=directional)

c. John is [**in front of**$_{s3}$] the tree.
spatial_signal(s3, cluster="in_front_of-1", semantic_type=directional)

`id`	me1, me2, me3, ...
`value`	number component
`unit`	measurement phrase component

Figure 8. *Attributes for MEASURE tag*

4.1.4. *MEASURE Tag*

The MEASURE tag is used to capture distances and dimensions for use in a measurement link. Its attributes are shown in Figure 8. Example (12) shows the annotation of two MEASUREs.

(12) The new tropical depression was about [**430 miles**$_{me1}$] ([**690 kilometers**$_{me2}$]) west of the southernmost Cape Verde Island, forecasters said.
measure(me1, value=430, unit=miles)
measure(me2, value=690, unit=kilometers)

Earlier versions of ISO-Space did not raise expressions such as these to first class objects in the annotation. It was felt, though, that trying to place measurement information into another tag, such as the MINK, would overload that tag. Using a separate MEASURE tag also allows for a more compositional annotation, which should allow for richer inferences down the road.

4.2. *ISO-Space Relationship Tags*

While ISO-Space 1.0 had one fairly overloaded relation tag, there are four distinct relation tags in the current specification:

(13) a. QSLINK – a qualitative spatial relationship between two locations;
b. OLINK – the orientation of a location or object relative to another;
c. MOVELINK – the representation of the path of an object in motion;
d. MLINK – the definition of the distance between two regions or the dimensions of a region.

Each of these tags is triggered by a specific kind of spatial element that was annotated in the text. QSLINKs are introduced by topological SPATIAL_SIGNALs, OLINKs by directional SPATIAL_SIGNALs, MOVELINK by MOTION events, and MLINK by MEASURE tags.

4.2.1. *Qualitative Spatial Link: QSLINK*

QSLINK is used in ISO-Space 1.4 to capture topological relationships between captured elements in the annotation. Essentially, it accounts for half of what the ver-

`id`	qsl1, qsl2, qsl3, ...
`relType`	{RCC8+}
`figure`	identifier of the place, path, spatial named entity, or event that is being related
`ground`	identifier of the place, path, spatial named entity, or event that is being related to
`trigger`	identifier of the spatial signal that triggered the link

Figure 9. *Attributes for QSLINK tag*

sion 1.0 QSLINK captured; the remaining information is captured by the OLINK in the new specification. The attributes of QSLINK are shown in Figure 9.

The `relType` attribute values come from a slightly extended set of RCC8 relations that was first used by SpatialML. The possible values include but are not limited to DC (disconnected), EC (external connection), and IN (disjunction of tangential and non-tangential proper part).

It is worth noting that while QSLINK is used exclusively for capturing topological relationships, which are only possible between two regions, the `figure` and `ground` attributes can accept IDs for both places and paths, which are more traditional regions, as well as spatial entities and events. In the latter cases, it is actually the region of space that is associated with the location of the entity or event that participates in the QSLINK. That is, the entity or event is *coerced* to a region for the purposes of interpreting this link.

In practice, a SPATIAL_SIGNAL with a `semantic_type` of topological introduces a QSLINK as shown in example (14).

(14) a. [**The book**$_{sne1}$] is [**on**$_{s1}$] [**the table**$_{sne2}$].
spatial_signal(s1, cluster="on-1", semantic_type=topological, directional)
qslink(qsl1, figure=sne1, ground=sne2, trigger=s1, relType=EC)

b. [**The light switch**$_{sne3}$] is [**on**$_{s2}$] [**the wall**$_{sne4}$].
spatial_signal(s1, cluster="on-2", semantic_type=topological, directional)
qslink(qsl2, figure=sne3, ground=sne4, trigger=s2, relType=PO)

4.2.2. *Orientation Link: OLINK*

Orientation links describe non-topological relationships between spatial objects. A SPATIAL_SIGNAL with a directional `semantic_type` triggers such a link. Rather than a simple relationship type, the OLINK is built around a specific frame of reference

id	ol1, ol2, ol3, ...
relType	NEAR, ABOVE, BELOW, FRONT, BEHIND, LEFT, RIGHT, NEXT TO, NORTH, ...
figure	identifier of the place, path, spatial named entity, or event that is being related
ground	identifier of the place, path, spatial named entity, or event that is being related to
trigger	identifier of the spatial signal that triggered the link
frame_type	ABSOLUTE, INTRINSIC, RELATIVE
referencePt	cardinal direction, ground entity, viewer entity
projective	TRUE, FALSE

Figure 10. *Attributes for OLINK tag*

type (Levinson, 2003) and a reference point. Figure 10 details the attributes for this link. [16]

The referencePt value depends on the frame_type of the link. Absolute OLINKs have a cardinal direction as a reference point. For intrinsic OLINKs, the reference point is the same identifier that was given in the ground attribute. For relative OLINKs, the ID for the viewer should be provided as the reference point. If the viewer is not explicit in the text, the special value "VIEWER" should be used. The projective attribute is a toggle that says whether the OLINK should have a projective interpretation. This information generally depends on what spatial signal triggered the OLINK. The examples in (15) show both projective and non-projective cases. Only the orientation links are shown.

(15) a. [**Boston**$_{pl1}$] is [**north of**$_{s1}$] [**New York City**$_{pl2}$].
olink(ol1, figure=pl1, ground=pl2, trigger=s1, relType="NORTH", frame_type=ABSOLUTE, referencePt=NORTH, projective=TRUE)

b. [**The dog**$_{sne1}$] is [**in front of**$_{s2}$] [**the couch**$_{sne2}$].
olink(ol2, figure=sne1, ground=sne2, trigger=s2, relType="FRONT", frame_type=INTRINSIC, referencePt=sne2, projective=FALSE)

c. [**The dog**$_{sne3}$] is [**next to**$_{s3}$] [**the tree**$_{sne4}$].
olink(ol3, figure=sne3, ground=sne4, trigger=s3, relType="NEXT TO", frame_type=RELATIVE, referencePt=VIEWER, projective=FALSE)

d. [**The hill**$_{pl3}$] is [**above**$_{s4}$] [**the town**$_{pl4}$].
olink(ol4, figure=pl3, ground=pl4, trigger=s4, relType="ABOVE", frame_type=INTRINSIC, referencePt=pl4, projective=TRUE)

16. In addition to frame-of-reference, values for cardinal directionality and orientation are provided. For further discussion of these issues, see Freksa and Zimmermann (1992) and Noyon *et al.* (2007).

e. [**The helicopter**$_{pl5}$] is [**above**$_{s5}$] [**the town**$_{pl6}$].
olink(ol4, figure=pl5, ground=pl6, trigger=s5, relType="ABOVE", frame_type=INTRINSIC, referencePt=pl4, projective=FALSE)

f. [**The book**$_{sne1}$] is [**on**$_{s1}$] [**the table**$_{sne2}$].
olink(ol4, figure=sne1, ground=sne2, trigger=s1, relType="ABOVE", frame_type=INTRINSIC, referencePt=sne2, projective=FALSE)

g. [**The light switch**$_{sne3}$] is [**on**$_{s2}$] [**the wall**$_{sne4}$].
olink(ol4, figure=sne3, ground=sne4, trigger=s2, relType="ABOVE", frame_type=INTRINSIC, referencePt=sne2, projective=FALSE)

4.2.3. *Movement Link: MOVELINK*

Movement links, which are introduced by motion events, capture information about an object in motion and the path a particular motion traverses. Following the exposition in Pustejovsky and Moszkowicz (2011), we assume the distinction made in Talmy (1985) between path constructions and manner constructions. Manner construction languages encode *path* information using directional prepositions (such as *to*, *from*, *towards*), particles (such as *out*, *away*, *up*), and other adjuncts, while the main (tensed) verb encodes the *manner-of-motion*. Path construction languages, on the other hand, encode *path* information in the main verb of the sentence, while adjunct Prepositional Phrases (PPs) optionally specify the *manner-of-motion*. In addition, we assume a modified classification of motion predicates, as defined in Muller (1998) and modified for annotation purposes in Pustejovsky and Moszkowicz (2008). It has the attributes shown in Figure 11. For example:

(16) a. [**John**$_{sne1}$] [**walked**$_{m1}$] from [**Boston**$_{pl1}$] to [**Cambridge**$_{pl2}$].
motion(id=m1, motion_type=MANNER, motion_class=MOVE)
movelink(mv1, trigger=m1, source=pl1, goal=pl2, mover=sne1, goal_reached=TRUE)

b. [**John**$_{sne1}$] [**arrived**$_{m1}$] in [**New York**$_{pl1}$].
motion(id=m1, motion_type=PATH, motion_class=REACH)
movelink(mv1, trigger=m1, goal=pl1, mover=sne1, goal_reached=TRUE)

4.2.4. *Metric Link: MLINK*

Metric relationships are captured with the MLINK tag. This tag can either describe the metric relationship between two spatial objects or the dimensions of a single object. The attributes are given in Figure 12.

When MLINK is used to describe an internal dimension of an object, the ID of the object should appear in the `figure` attribute. The annotator may either repeat the ID in the `ground` attribute or leave this attribute out. The examples below show several ways in which MLINK is used. Examples (17c) and (17d) show the unique case when a stative path, or a path that does not involve traversal, is used to describe the dimensions

`id`	mvl1, mvl2, mvl3, ...
`trigger`	identifier of the motion event that triggered the link
`source`	identifier of the place, path, spatial named entity, or event at the beginning of the path
`goal`	identifier of the place, path, spatial named entity, or event at the end of the path
`mover`	identifier of the entity that moves along the path
`goal_reached`	TRUE, FALSE
`pathID`	identifier of a path that is equivalent to the one described by the MOVELINK

Figure 11. *Attributes for MOVELINK tag*

`id`	ml1, ml2, ml3, ...
`figure`	identifier of a spatial object
`ground`	identifier of the related spatial object, if there is one
`relType`	DISTANCE, LENGTH, WIDTH, HEIGHT, GENERAL_DIMENSION
`val`	NEAR, FAR, identifier of a measure
`endPoint1`	identifier of a spatial object at one end of a stative path
`endPoint2`	identifier of a spatial object at the other end of a stative path

Figure 12. *Attributes for MLINK tag*

of a location. In such a case, the optional attributes `endPoint1` and `endPoint2` are used.

(17) a. The new [**tropical depression**$_{sne1}$] was about [**430 miles**$_{me1}$] ([**690 kilometers**$_{me2}$]) west of the [**southernmost Cape Verde Island**$_{pl1}$], forecasters said.
mlink(ml1, relType = DISTANCE, figure=sne1, ground=pl1, val=me1)

b. [**The football field**$_{sne2}$] is [**100 yards**$_{me2}$] long.
mlink(ml2, relType = LENGTH, figure=sne2, ground=sne2, val=me2)

c. [**Times Square**$_{pl2}$] stretches from [**42nd**$_{p1}$] to [**47th streets**$_{p2}$].
mlink(ml3, relType = GENERAL_DIMENSION, figure=pl2, ground=pl2, endPoint1=p1, endPoint2=p2)

d. [**The office**$_{pl3}$] stretches for [**25 feet**$_{me3}$] from [**the bookcase**$_{sne3}$] to [**the white board**$_{sne4}$].
mlink(ml4, relType=GENERAL_DIMENSION, figure=pl4, ground=pl3, val=me3, endPoint1=sne3, endPoint2=sne4)

e. [**The hot dog stand**$_{sne5}$] near [**Macy's**$_{sne6}$].
mlink(ml5, relType=GENERAL_DIMENSION, figure=sne5, ground=sne6, val=NEAR)

5. Outstanding Issues

The above specification leaves several issues unanswered concerning the representation of spatial information as a specification language. Perhaps the most significant is the absence of a *native* representation of the 3D objects denoted by linguistic expressions, along with the associated functions we naturally ascribe to them. For example, the topological relationship between a glass and the liquid it holds should convey more than the RCC8 relations of EC or TPP, neither of which is exactly correct. Rather, within a 3D interpretation, the appropriate relation should express containment of a region within a convex volume. Similarly, the interpretation of an object inside a box should also make reference to such a containment relation, rather than a mere EC value. That is, more credence should be given to image-schematic accounts of spatial categories and how this impacts the spatial configurational relation that are denoted by real-world spatial situations (Frank and Raubal, 1999; Kuhn, 2007). This follows the methodology adopted in Mani and Pustejovsky (2012) to develop richer specifications incorporating broader coverage of the language data, in a layered fashion. These issues are currently being examined within the ISO-Space working group for inclusion into the specification language, as it develops.

Acknowledgements

We would like to thank Inderjeet Mani for his contribution to the research and development of the specification. We would also like to thanks the members of the ISO-Space Working Group for their significant input to the current specification as well as the participants of the Brandeis and Airlie Workshops. Finally, we gratefully acknowledge the extremely helpful comments from two reviewers for the Journal. Part of this research was funded under a NURI grant HM1582-08-1-0018 by the National Geospatial Agency. All errors and mistakes are, of course, the responsibilities of the authors.

6. References

Asher N., Vieu L., "Towards a geometry of common sense: a semantics and a complete axiomatisation of merotopology", *Proceedings of IJCAI95*, Montreal, Canada, 1995.

Bateman J. A., Hois J., Ross R., Tenbrink T., "A linguistic ontology of space for natural language processing", *Artificial Intelligence*, 2010.

Blaylock N., "Semantic annotation of street-level geospatial entities", *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on*, IEEE, p. 444-448, 2011.

Bunt H., "A methodology for designing semantic annotation languages exploiting syntactic-semantic iso-morphisms", *Proceedings of ICGL 2010, Second International Conference on Global Interoperability for Language Resources*, 2010.

Cohn A. G., Renz J., "Qualitative spatial representation and reasoning", , vol. 46, p. 1-2, 2001.

Denis M., “The description of routes: a cognitive approach to the production of spatial discourse”, *Current Psychology of Cognition*, vol. 4(16), p. 409-458, 1997.

Denis P., Muller P., “Predicting globally-coherent temporal structures from texts via endpoint inference and graph decomposition”, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, 2011.

Duckham M. Winter S. R. M., “Including landmarks in routing instructions”, *Journal of Location Based Services*, vol. 4(1), p. 28-52, 2010.

Edwardes A., Purves R., Bircher S., Matyas C., “Deliverable 1.4: Concept ontology experimental report. , TRIPOD Project”, , Technical report, Sheffield University, 2007.

Egenhofer M., Franzosa R., “Point-set topological spatial relations”, *International Journal of Geographical Information Science*, vol. 5, n° 2, p. 161-174, 1991.

Frank A. U., Raubal M., “Formal specification of image schemata – a step towards interoperability in geographic information systems”, *Spatial Cognition and Computation*, vol. 1, p. 67-101, 1999. 10.1023/A:1010004718073.

Freksa C., “Using orientation representation for qualitative spatial reasoning”, *in* A. Frank, I. Campari, U. Formentini (eds), *Theories and Methods of Spatio- temporal Reasoning in Geographic Space: Proceedings of the International Conference GIS - From Space to Territory*, Pisa, Italy, p. 162-178, 1992.

Freksa C., Zimmermann K., “On the utilization of spatial structures for cognitively plausible and efficient reasoning”, *Proc. of the Conf. on Systems, Man, and Cybernetics*, Chicago, USA, p. 261-266, 1992.

Galton A., “Towards a qualitative theory of movement”, *in* A. U. Frank, W. Kuhn (eds), *Spatial Information Theory: A Theoretical Basis for GIS (Proceedings of International Conference COSIT’95)*, Springer-Verlag, Semmering, Austria, p. 377-396, September, 1995.

Ide N., Romary L., “International standard for a linguistic annotation framework”, *Natural Language Engineering*, vol. 10, n° 3-4, p. 211-225, 2004.

Klippel A., Tappe H., Kulik L., Lee P., “Wayfinding choremes: a language for modeling conceptual route knowledge”, *Journal of Visual Languages and Computing*, vol. 16, n° 4, p. 311-329, 2005.

Kordjamshidi P., Moens M., van Otterlo M., “Spatial role labeling: task definition and annotation scheme”, *Proceedings of LREC 2010 - The sevent international conference on langauge resources and evaluation*, 2010.

Kuhn W., “An image-schematic account of spatial categories”, *Lecture Notes in Computer Science 4736*, p. 152-168, 2007.

Kurata Y., Egenhofer M., “The 9+ intersection for topological relations between a directed line segment and a region”, *in* B. Gottfried (ed.), *Workshop on Behaviour and Monitoring Interpretation*, Germany, p. 62-76, September, 2007.

Levinson S., *Space in Language and Cognition: Explorations in Cognitive Diversity*, Language, culture, and cognition, Cambridge University Press, 2003.

Levit M., Roy D., “Interpretation of spatial language in a map navigation task”, *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 2007.

Ligozat G., “Reasoning about cardinal directions”, *Journal of visual languages and computing*, vol. 9, p. 23-44, 1998.

Mani I., Doran C., Harris D., Hitzeman J., Quimby R., Richer J., Wellner B., Mardis S., Clancy S., "SpatialML: annotation scheme, resources, and evaluation", *Language Resources and Evaluation*, vol. 44, n° 3, p. 263-280, September, 2010.

Mani I., Hitzeman J., Richer J., Harris D., Quimby R., Wellner B., "SpatialML: annotation scheme, corpora, and tools", *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, European Language Resources Association (ELRA), Marrakech, Morocco, may, 2008.

Mani I., Pustejovsky J., *Interpreting Motion: Grounded Representations for Spatial Language*, Oxford University Press, 2012.

Mani I., Wilson G., "Robust temporal processing of news", *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL2000)*, New Brunswick, New Jersey, p. 69-76, 2000.

Mitra D., "Modeling and reasoning with star calculus: an extended abstract", *Eighth International Symposium on AI and Mathematics*, 2004.

Moszkowicz J. L., Dynamic Location Updates: The Impact of Motion on Discourse Structure, PhD thesis, Brandeis University, 2011.

Moszkowicz J. L., Pustejovsky J., "ISO-Space: towards a spatial annotation framework for natural language", *Processing Romanian in Multilingual, Interoperational and Scalable Environments*, 2010.

Muller P., "A qualitative theory of motion based on spatio-temporal primitives", *in* A. G. Cohn, L. Schubert, S. C. Shapiro (eds), *KR'98: Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, San Francisco, California, p. 131-141, 1998.

Muller P., Prévot L., "Grounding information in route explanation dialogues.", *in* T. T. K. R. Coventry, J. Bateman (eds), *Spatial Language and Dialogue*, Oxford University Press, 2009.

Noyon V., Claramunt C., Devogele T., "A relative representation of trajectories in geographical spaces", *GeoInformatica*, 2007.

Prévot L., *Structures sémantiques et pragmatiques pour la modélisation de la cohérence dans des dialogues finalisés*, intelligence artificielle : Toulouse 3 : 2004, 2004.

Pustejovsky J., Castano J., Ingria R., Saurí R., Gaizauskas R., Setzer A., Katz G., "TimeML: robust specification of event and temporal expressions in text", *IWCS-5, Fifth International Workshop on Computational Semantics*, 2003. www.timeml.org.

Pustejovsky J., Knippen R., Littman J., Saurí R., "Temporal and event information in natural language text", *Language Resources and Evaluation*, vol. 39, p. 123-164, May, 2005.

Pustejovsky J., Moszkowicz J., "The qualitative spatial dynamics of motion", *The Journal of Spatial Cognition and Computation*, 2011.

Pustejovsky J., Moszkowicz J. L., "Integrating motion predicate classes with spatial and temporal annotations", *Proceedings of COLING 2008*, Manchester, UK, 2008.

Pustejovsky J., Moszkowicz J. L., Verhagen M., "ISO-Space: the annotation of spatial information in language", *Proceedings of ISA-6: ACL-ISO International Workshop on Semantic Annotation*, Oxford, England, January, 2011.

Randell D., Cui Z., Cohn A., "A spatial logic based on regions and connections", *in* M. Kaufmann (ed.), *Proceedings of the 3rd Internation Conference on Knowledge Representation and REasoning*, San Mateo, p. 165-176, 1992.

Raubal M., Winter S., “Enriching wayfinding instructions with local landmarks”, Springer, p. 243-259, 2002.

Renz J., Mitra D., “Qualitative direction calculi with arbitrary granularity”, *Proceedings of the 8th Pacific Rim International Conference on Artificial Intelligence*, p. 65-74, 2004.

Setzer A., Temporal Information in Newswire Articles: An Annotation Scheme and Corpus Study, PhD thesis, University of Sheffield, UK, 2001.

Talmy L., “Lexicalization patterns: semantic structure in lexical forms”, *in* T. Shopen (ed.), *Language typology and semantic description Volume 3: Grammatical categories and the lexicon*, Cambridge University Press, p. 36-149, 1985.

Taylor H., Tversky B., “Spatial mental models derived from survey and route descriptions”, *Journal of Memory and Language*, vol. 31, p. 261-292, 1992.

Tversky B., Lee P. U., “Pictorial and verbal tools for conveying routes”, Springer, p. 51-64, 1999.

Vogel A., Jurafsky D., “Learning to follow navigational directions”, *Proceedings of ACL-2010*, Uppsala, Sweden, 2010.

Wing B., Baldridge J., “Simple supervised document geolocation with geodesic grids”, *Proceedings of ACL*, p. 955-964, 2011.