Using ISO-Space for Annotating Spatial Information

James Pustejovsky, Jessica L. Moszkowicz, and Marc Verhagen
{jamesp,jlittman,marc}@cs.brandeis.edu

Computer Science Department, Brandeis University, Waltham, MA 02454 USA

Abstract. In this paper, we describe ISO-Space, an annotation language currently under community development for encoding spatial and spatiotemporal information as expressed in natural language text. After reviewing the requirements of a specification for capturing such knowledge from linguistic descriptions, we demonstrate how ISO-Space aims to address these problems. ISO-Space is an emerging resource that is still in its early stages of development; hence community involvement from multiple disciplines and potential users and consumers is necessary, if it is to achieve a level of descriptive adequacy and subsequent adoption. We describe the genres of text that are being used in a pilot annotation study, in order to both refine and enrich the specification language.

1 Motivation and Problem Definition

Human languages impose diverse linguistic constructions for expressing concepts of space, of spatially-anchored events, and of spatial configurations that relate in complex ways to the situations in which they are used. One area that deserves further development regarding the connection between natural language and formal representations of space is the automatic enrichment of textual data with spatial annotations. There is a growing demand for such annotated data, particularly in the context of the semantic web. Moreover, textual data routinely make reference to objects moving through space over time. Integrating such information derived from textual sources into a geosensor data system can enhance the overall spatiotemporal representation in changing and evolving situations, such as when tracking objects through space with limited image data. Hence, verbal subjective descriptions of spatial relations need to be translated into metrically meaningful positional information. A central research question currently hindering progress in interpreting textual data is the lack of a clear separation of the information that can be derived directly from linguistic interpretation and further information that requires contextual interpretation. Markup schemes should avoid over-annotating the text, in order to avoid building incorrect deductions into the annotations themselves. Solutions to the language-space mapping problem and its grounding in geospatial data are urgently required for this.

In fact, establishing tighter formal specifications of the relationship between language and space has proved to be a considerable challenge. In this paper, we introduce an annotation framework called ISO-Space that aims to be just such a specification. ISO-Space incorporates the annotations of static spatial information, borrowing from the SpatialML scheme [21, 19], along with a new markup language called Spatiotemporal Markup Language (STML) [25] that focuses on locating events in space. The term "ISO-Space" is adopted as this effort is being developed within the ISO TC37/SC4 technical subcommittee on language resource management as part six of the Semantic Annotation Framework, where the goal is to create a new standard for capturing spatial and spatiotemporal information.

There are many applications and tasks which would benefit from a robust spatial markup language such as ISO-Space. These include, among others:

(1) a. Creating a visualization of objects from a verbal description of a scene;
  b. Identifying the spatial relations associated with a sequence of processes and events from a news article;
  c. Determining an object location or tracking a moving object from a verbal description;
  d. Translating viewer-centric verbal descriptions into other relative descriptions or absolute coordinate descriptions.
  e. Constructing a route given a route description.
  f. Constructing a spatial model of an interior or exterior space given a verbal description.
  g. Integrating spatial descriptions with information from other media.

The goal of ISO-Space is not to provide a formalism that fully represents the complexity of spatial language, but rather to capture these complex constructions in text to provide an inventory of how spatial information is presented in natural language.

To illustrate the application of the proposed annotation language, let us consider four scenarios in greater detail. Imagine first a dynamically evolving situation that is being witnessed by multiple viewers, each reporting with multiple feeds, such as Twitter or SMS texts, possibly with geolocated and time-stamped messaging capabilities. The issues of spatial representation and reasoning from language in such a scenario are quite complex, involving 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. A spatial annotation language, to be useful for such cases, must minimally capture these dimensions of spatial information, and allow for this data to supplement or complement information coming from other modalities representing contextual or spatial information.

A related scenario involves the integration of geosensor data over a monitored area with viewer feeds. 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 represented twice in the scene, once at time t1 and once at time t2. The only other object in the scene is a newsstand.

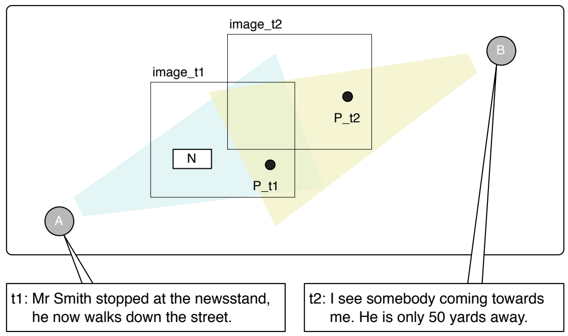


Fig. 1. Multiple-viewer Object Tracking

A second scenario that highlights the demands on a spatial annotation language involves the identification and interpretation of locations and objects in motion as reported in a news article, such as the one below.

> NYT: July 15, 2010
> CAIRO — A Libyan ship that tried to break Israel's blockade of Gaza docked in the Egyptian port of El Arish on Thursday afternoon as the ship's sponsor [...] said that the boat had shifted course because the Israeli government agreed to allow Libya to support building and reconstruction in Gaza.

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. TimeML ([23]) was designed with just such applications in mind. Extending this paradigm to space, SpatialML([19]) has proved fairly robust in geolocating geographic entities and facilities in text. The challenge presented by texts such as this, however, is that the motion, the moving objects, and the paths along which these motions occur, must be identified in order to understand the spatial consequences of the text. Here the problem involves tracking the movement of the ship from a region identified as part of the blockade near Gaza to a location in Egypt identified as a port.

The third scenario involves 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.

The issues raised by this example primarily involve the interpretation of orientation, as well as the fact that the viewpoint is being updated as the directions are traversed by the reader.

Finally, the fourth scenario involves an application involving object labeling and image annotation, where the goal is to not only label the objects, but explicitly identify their spatial relations to each other, both intrinsic and relative, from the viewer's perspective. This involves taking viewer-centric descriptions of objects relative to landmarks (with GPS locations on the source) and converting these from relative descriptions to normalized or absolute frame or reference representations. In the image below, for example, an annotator might label the objects as shown, but also the relations: the red car is in front of the Hall; the Prudential is to the left of the Christian Science Monitor; the truck is in the street, and so on.

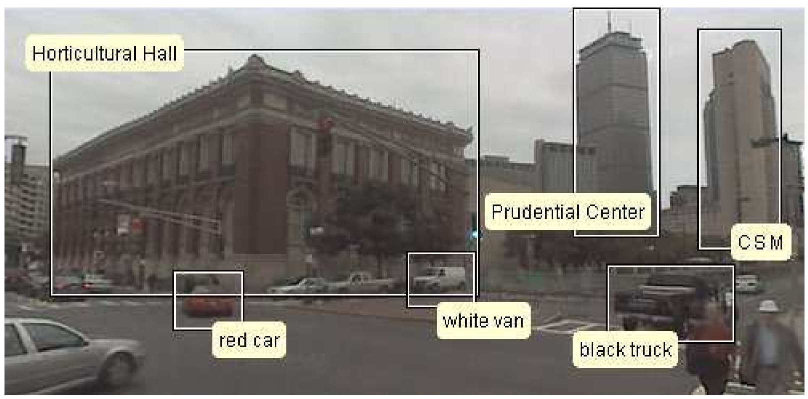


Fig. 2. Image Labeling

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 gazateers.

In the discussion below, we first 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.

2 Requirements

ISO-Space is designed to capture two kinds of information: spatial information and spatiotemporal information. While still in development, it is clear that the conceptual inventory for spatial language annotation must include (at least) the following notions:

(2) a. Locations (regions, spatial objects): Geographic, Geopolitical Places, Functional Locations.
b. Entities viewed as Spatial Objects.
c. Paths as Objects: routes, lines, turns, arcs
d. Topological relations: *inside*, *connected*, *disconnected*.
e. Direction and Orientation: *North*, *downstream*.
f. Time and space measurements: units and quantities for spatial and temporal concepts.
g. Object properties: intrinsic orientation, dimensionality, size, shape.
h. Frames of reference: absolute, intrinsic, relative.
i. Spatial Functions: *behind the building*, *twenty miles from Boulder*.
j. Motion: tracking moving objects over time.

We will refer to constructions that make explicit reference to the spatial attributes of an object or spatial relations between objects as *spatial expressions*. Linguists traditionally divide spatial expressions into at least four grammatically defined classes:

(3) a. Spatial Prepositions and Particles: *on*, *in*, *under*, *over*, *up*, *down*, *left of*;
b. Verbs of position and movement: *lean over*, *sit*, *run*, *swim*, *arrive*;
c. Spatial attributes: *tall*, *long*, *wide*, *deep*;
d. Spatial Nominals: *area*, *room*, *center*, *corner*, *front*, *hallway*.

Unlike the fairly well-behaved list of 13 values for temporal relations in language (as encoded in ISO-TimeML), spatial prepositions are notoriously ambiguous and context dependent. Not only are there vastly more configurations possible between objects construed as spatial regions, but languages are idiosyncratic in how spatial information is encoded through different linguistic expressions. For this reason, we will have to define constraints that allow for underspecified semantic interpretations for several of the concepts introduced in our abstract syntax. These will need to communicate with various lexical [20, 8, 16] and spatial ontological resources [9, 3], in order to help disambiguate and more fully determine the semantics of relation types from the specification [12].

There has been considerable research on the linguistic behavior of spatial predicates and prepositions in language [28, 15, 14, 5, 6]. Within qualitative spatial reasoning (QSR), work has recently started to focus on incorporating mereotopological concepts into the calculus of relations between regions [27, 13, 7, 18].

The focus of SpatialML is to mark up spatial locations mentioned in texts while allowing integration with resources that provide information about a given domain, such as physical feature databases and gazetteers. The core SpatialML tag is the PLACE tag, which includes attributes `type` (country, continent, populated place, building, etc.), `gazref` (a reference to a gazetteer entry), and `LatLong` (a latitude and longitude pair). Complex locations such as *Pacific coast of Australia* and *the hot dog stand behind Macy's* are annotated using the LINK and RLINK tags, respectively, encoding topological and relative relations between places. While SpatialML is one of the cornerstones of ISO-Space, it needs to be extended to account for more of the complexity of spatial language. Most notably, SpatialML was not designed to capture implicit places such as the one described by the phrase *behind Macy's* rather than the more complete example *the hot dog stand behind Macy's*.

In addition to capturing implicit spatial information, ISO-Space will include additional properties of locations such as orientation and metric relations between objects, the shape of an object, the size of an object, elevation, granularity, aggregates and distributed objects, and objects in motion. A major focus of the ISO-Space effort is to encode as complete a range of verbal descriptions of spatial properties as possible. Spatial relations as expressed in language express aspects of (at least) the five parameters of meaning listed below:

(4) a. Topological: *in*, *inside*, *touching*, *outside*;
 b. Orientational (with frame of reference): *behind*, *left of*, *north of*;
 c. Metric: *20 miles away*, *near*, *close by*;
 d. Topological-orientational: *on*, *over*, *below*;

The other components of ISO-Space, STML and TimeML/ISO-TimeML [23, 24], bring in the spatio-temporal dimension. This dimension is captured by introducing spatial events, most specifically by annotating motions. The Spatio-Temporal Markup Language proposes a set of motion classes based on the classifications by [22] and refined in [25]. The latter work performed a comparison of Muller's classes as well as those from Vieu [30] and Asher and Sablayrolles [1] in order to arrive at the most expressive set of motion predicates. A mapping to existing resources such as FrameNet [2] and VerbNet [16] was also included[1]. The results of the analysis are summarized in Table 1.

Category	FrameNet	Muller	Asher/Sablayrolles
Move (*run, fly, drive*)	Motion, Self_motion	no mapping	no mapping
Move_External (*traverse, pass*)	Traversing	External	no mapping
Move_Internal (*walk around*)	Motion	Internal	no mapping
Leave (*leave, desert*)	Departing	Leave	*partir, sortir*
Reach (*arrive, enter, reach*)	Arriving	Reach	*arriver, enter*
Detach (*take off, pull away*)	no mapping	no mapping	*décoller*
Hit (*land, hit*)	Impact	Hit	*se poser*
Follow (*follow chase*)	Co-theme	no mapping	no mapping
Deviate (*flee, run from*)	Fleeing	no mapping	*dévier*
Stay (*remain, stay*)	State_continue	no mapping	no mapping

Table 1. The ten motion classes from STML

The meanings for each of these classes will correspond to a semantic interpretation of motion concepts specified in the abstract syntax, as mentioned above.

[1] The mapping to VerbNet proved not to be useful at the time. For more information about the encoding of change of location in VerbNet, see [31].

3 Description of the ISO-Space Elements

We now informally describe the ISO-Space elements, using example annotations from real text when possible. Our goal is to present a glimpse of the richness required to capture spatial information in language. Further details can be found in [26]. There are two major types of elements:

1. ENTITIES: location, spatial_entity, motion, state, event_path, path;
2. SPATIAL RELATIONS: topological, relational, and distance.

Along with these two main classes, there are some minor elements, such as spatial signals. In the discussion that follows, we explain these elements in more detail, beginning with the entities.

Location. An *inherently grounded spatial* entity, a location includes geospatial entities such as mountains, cities, rivers, etc, as well as administrative entities like towns and counties. It also includes classificatory and ontological spatial terms, such as *edge*, *corner*, *intersection*, and so forth.

Some examples of locations are presented in (5) and (6). Note that in (6), the string *430 miles (690 kilometers) west of the Sounthernmost Cape Verde Island* is not annotated as one composite location that embeds another location. The distance and direction are not included in the location, but expressed as signals. We will return to this example later.

(5) A Libyan ship that tried to break Israels blockade of [**Gaza**$_{loc1}$] docked in the [**Egyptian port of El Arish**$_{loc2}$] on Thursday afternoon.

(6) The new tropical depression was about 430 miles (690 kilometers) west of the [**southernmost Cape Verde Island**$_{loc3}$], forecasters said.

The attributes for the LOCATION tag are largely inherited from SpatialML's PLACE element [19] and include:

id	loc1, loc2, loc3, ...
type	BODYOFWATER, CELESTIAL, CIVIL, CONTINENT, COUNTRY, GRID, LATLONG, MTN, MTS, POSTALCODE, POSTBOX, PPL, PPLA, PPLC, RGN, ROAD, STATE, UTM
continent	AF, AN, AI, AU, EU, GO, LA, NA, PA, SA
country	a two letter ISO 3166 country code
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
latLong	a coordinate from the gazetteer
mod	a spatially relevant modifier
dcl	true or false

The *Document Creation Location* or `DCL` is a special location that serves as the "narrative location"[2]. If the document includes a DCL, it is generally specified at the beginning of the text. For example, a news article may explicitly state where the story is being written. For Twitter-like data, one can imagine that the entry may be explicitly geolocated or have some explicit mention of a location from which the entry was made.

We are pragmatic on whether there should be a distinction between place, location, and region. Region, as a geometric concept, is currently not a first-class citizen in ISO-Space. The LOCATION element covers both locations and places (where a place is considered a functional category), and is assumed to be associated with a region whenever appropriate. We do acknowledge however that, in some cases locations, are mapped to lines instead of regions. If it becomes apparent that this distinction is relevant to the annotation, then the specification can be modified accordingly.

Spatial_entity. An entity that is not inherently a LOCATION, but one which is identified as participating in a spatial relation is tagged as a SPATIAL_ENTITY. It can be an entity such as a *car* or *building* or an individual like *John*. These can also be event-like entities like *traffic jam* or *hurricane*. A SPATIAL_ENTITY is only annotated in the context of an explicit spatial relation. Each SPATIAL_ENTITY inherently defines a location and can be the location for other spatial entities, as in *John is in the car*.

With this element, there is a top-level distinction between spatial entities and the locations they are in. In many cases, a spatial entity will act as a location for other objects. This raises the issue of whether entities like *building* in *The statue is in the building* are annotated as locations or spatial entities. We resolve this by stipulating that these entities are never annotated as locations but always as spatial entities, even in a case like *the president is in the building*.

Below we repeat examples (5) and (6), with annotations for spatial entities added.

(7) A [**Libyan ship**$_{se1}$] that tried to break Israels [**blockade**$_{se2}$] of [**Gaza**$_{loc1}$] docked in the [**Egyptian port of El Arish**$_{loc2}$] on Thursday afternoon.

(8) The new [**tropical depression**$_{se3}$] was about 430 miles (690 kilometers) west of the [**southernmost Cape Verde Island**$_{loc3}$], forecasters said.

Spatial entities share some attributes with locations. Like LOCATION, SPATIAL_ENTITY has a `type` attribute with the following possible values: FAC, VEHICLE, PERSON, and DYNAMIC EVENT. These values are preliminary and additional values could be derived from existing ontologies.

id	se1, se2, se3, ...
type	FAC, VEHICLE, PERSON, DYNAMIC EVENT, ...
latLong	a coordinate
mod	a spatially relevant modifier

[2] The DCL is based on ISO-TimeML's Document Creation Time (DCT).

Motion. A MOTION is an *inherently spatial* event, involving a change of location. Motions include verbs like *leave*, *move*, and *approach*, as in the example below.

(9) ... it could become a Category 2 hurricane as [**it**$_{se5}$] [**approaches**$_{m2}$] the [**Caribbean**$_{loc3}$]

Motions have attributes for motion type, motion class, mode, speed and event path identifier.

id	m1, m2, m3, ...
motion_type	MANNER or PATH
motion_class	MOVE, MOVE_EXTERNAL, MOVE_INTERNAL, LEAVE, REACH, DETACH, HIT, FOLLOW, DEVIATE, STAY
mode	either the mode of transportation or a reference to an additional motion of type manner
speed	
event_pathID	ep1, ep2, ep3, ...

The `motion_type` attribute refers to the two distinct strategies for expressing concepts of motion in language: *path constructions* and *manner-of-motion constructions* [29]. 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]$_p$.

While some predicates are clearly path or manner-of-motion examples, it is often the case, as seen in the above examples that both a path and manner construction are composed. In such a case, the mode attribute is used. This attribute is used exclusively with motions of type path. When the mode of transportation is specified as in (10a), this is represented in the mode attribute. In some instances, the manner will be given as a second motion as in *John left the store running.* Here, both *left* and *running* are tagged as motions and the ID for the latter is given as the mode value for the former.

As mentioned earlier, motion classes are taken from [25], which was based on the motion classes in [22]. These classes are associated with a spatial event structure that specifies, amongst others, the spatial relations between the arguments of the motion verb at different phases of the event. This is a preliminary classification and does, for example, not deal properly with the *approach* example in (9) which does not fit in any of the ten classes. The `speed` attribute is included because it allows for spatio-temporal inferences. Note however that speed may be better analyzed as a separate tag similar to the DISTANCE element with `unit` and `measure` attributes (described later in this section). Motions generate and participate in EVENT_PATH elements. Some characteristics of motions, like direction, are expressed in this additional tag.

State. A STATE, or, perhaps, a SPATIAL_STATE, is a situation that does not involve a change of location, but one which is identified as participating in a spatial relation[3]. In the sentence below, *slept* is a state that is spatially related to a location.

(11) ... which meant we [**slept**$_{e1}$] on the crowded [**second floor**$_{loc1}$]

Path. A PATH is a functional location where the focus is on the potential for traversal. It includes common nouns like *road* and *river* and proper names like *Route 66* and *Kangamangus Highway*. The distinction between locations and paths is not always clear-cut. Take for example the case of *river* in *follow the river* and *cross the river*. The first case clearly is a path (assuming we are on a boat), but one could make a case that in the second case the traversal functionality is not accessed and therefore *river* should be annotated as a location. Whether annotated as a location or a path, the inferences that can be made are identical and we propose to leave decisions regarding these cases to project-specific annotation guidelines.

Traversals of paths are inherently directional but paths themselves are not. Paths typically have begin points and end points (although these are often not expressed in the text), but the choice of what is the begin point and what the end point is often arbitrary.

The attributes of PATH are a subset of the attributes of the LOCATION element, but with the beginID and endID elements added:

id	p1, p2, p3, ...
type	BODYOFWATER, MTS, ROAD
beginID	identifier of a location tag
endID	identifier of a location tag
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

Event_path. The implicit path that is introduced by virtue of a motion event is captured with the EVENT_PATH tag. It is defined by the motion event as well as by elements introduced in the immediate context of the motion event, like begin point, end point, path region, and intermediate points. Event paths differ from paths in three respects: (i) they are always directed even though the direction may not have been made explicit in the text, (ii) the actual physical path followed does not have to be specified, and (iii) they are always associated with a motion. In example (12), the *moving* motion creates an event path with an explicitly specified direction.

[3] Note that this is different from other treatments such as [4, 11, 10].

(12) The [**depression**$_{se1}$] was [**moving**$_{m1}$] [**westward**$_{sig1}$] at about [**17 mph**$_{sig2}$] ([**28 kph**$_{sig3}$]) and was expected to [**continue**$_{e1}$] that motion for the next day or two.
Motion(m1, speed=17mph,28kph)
EventPath(ep1, source=m1, direction=WEST, moving_object=s1, signals=[sig1])

The `sourceID` attribute holds the ID of the motion event. When possible the location IDs for the endpoints and a spatial signal ID, if applicable, are given in the EVENT_PATH element, which includes the following attributes:

id	ep1, ep2, ep3, ...
sourceID	a motion identifier
objectID	identifier of the moving object, typically a spatial entity or an event
beginID	identifier of the location, spatial entity or event that is at the beginning of the path
endID	identifier of the location, spatial entity or event that is at the end of the path
end_reached	true or false
midIDs	list of identifiers of midpoints
pathIDs	list of identifiers of a paths
direction	NORTH, WEST, SOUTH, ...
directionID	the id of the signal that specified the direction
length	the distance between the begin and end point
lengthID	identifier of the signal that specified the length

All of these attributes, except for `id` and `sourceID` can be empty. The `end_reached` boolean is used to distinguish between cases like *John left for Boston* and *John arrived in Bosten*, where Boston was reached in the latter case but not necessarily in the former. The `pathIDs` attribute is motivated by examples such as *He took I-95 towards New York*, where the event path refers to a path, but where that path is not to be equated with the event path or part of the event path.

The remaining ISO-Space tags capture information about spatial relations between the entities introduced above.

Qualitative spatial link. A *qualitative spatial link* details the spatial relationship between two spatial objects (figure and ground)[4], where a spatial object can be a location, a path, a spatial entity, or a state. Typically, one of the objects is a location, but this is not necessarily the case.

There are two kinds of spatial relations: *topological spatial relations* and *relational spatial relations*. A topological QSLINK introduces a relation type from

[4] The terms *figure* and *ground* are used primarily for the benefit of relational links. For topological links, these simply serve as labels for the first and second arguments in the link, respectively.

the extended RCC8 set which includes all RCC8 relations as well as the IN relation introduced by SpatialML, which is a disjunction of the RCC8 relations that imply that one region is inside the other:

DC	Disconnected
EC	External Connection
PO	Partial Overlap
EQ	Equal
TPP, TPPi	Tangential Proper Part and its inverse
NTTP, NTTPi	Non-Tangential Proper Part and its inverse
IN	disjunction of TTP and NTTP

A relational QSLINK is a spatial link that essentially defines one region in terms of another, as in *the car in front of the house* where the position of the car is specified relative to the position of the house. These spatial links, as well as the topological links, are generally introduced by a spatial signal like *in front of*. Relational QSLINKS get a relation from one or more closed-class sets with spatial markers. Typically, these markers are normalized versions of function words or spatial measures in the text. For example, if the function word is *northwest of*, then the `relation` for the QSLINK will be NORTHWEST.

ISO-Space will include an inventory of relation type values that each have a well-defined semantics. This inventory will include directions like NORTH and SOUTHEAST as well as sense inventories for prepositions where we will introduce values like ON_1 and ON_2 that could account for the differences between *the vase on the table* and *the fly on the wall*. Note that, in a sense, the topological relations are a subpart of this inventory. We intend to provide a core inventory that can be expanded at the discretion of individual applications. A sub inventory could be specific to the objects that are being related. For example, we could imagine an inventory of relations from Allen's time interval algebra (before, overlap, etcetera) for spatial relations between paths since paths are one-dimensional.

A QSLINK element has the following attributes:

id	qsl1, qsl2, qsl3, ...
type	TOPOLOGICAL or RELATIONAL
relation	{RCC8+} ∪ {CORE_INVENTORY}
figureID	identifier of the location, spatial entity, or state that is being related
groundID	identifier of the location, spatial entity, or state that is being related to
rframe	ABSOLUTE (default) or RELATIVE or INTRINSIC
viewer	identifier of the viewer for the reference frame
spatial_signalID	identifier of the spatial signal

Example (13) includes two states that are both related to the location. The value for the relation type is IN as suggested by the spatial signal *in*.

(13) ... the Israeli government agreed to allow Libya to support [**building**$_{e1}$] and [**reconstruction**$_{e2}$] [**in**$_{sig1}$] [**Gaza**$_{loc1}$]
QSLink(qsl1, figureID=e1, groundID=loc1, type=topological, relation=IN)
QSLink(qsl2, figureID=e2, groundID=loc1, type=topological, relation=IN)

In (14), a spatial entity is related to a location. Both a topological and a relative spatial link are added, which is perfectly fine since the two relations are not contradictory. However, guidelines for annotation may stipulate that when a spatial link like qsl1 is added, then there is no need to add qsl2 since the EXTERNAL relation can be derived from the semantics of WEST.

(14) The new [**tropical depression**$_{se1}$] was about [**430 miles (690 kilometers)**$_{sig1}$] [**west of**$_{sig2}$] the [**southernmost Cape Verde Island**$_{loc1}$], forecasters said.
QSLink(qsl1, figureID=se1, groundID=loc1, type=relational, relation=WEST, spatial_signalID=sig2)
QSLink(qsl2, figureID=se1, groundID=loc1, type=topological, relation=EC)

When a spatial link is labeled as relational, this not only requires that the relation type come from the core inventory, but also that the `rframe` attribute is used. Be default, all qualitative spatial links are assumed to have an ABSOLUTE frame. The INTRINSIC frame can be used when the interpretation of the link relies on an intrinsic property of the ground element in the link. For example, absent of any other context, the example *John sat in front of the computer* produces a relative spatial link with an INTRINSIC `rframe`.

A RELATIVE value is used in `rframe` when the link is to be interpreted from a particular viewer's vantage point. Example (15) shows a possible annotation of a sentence given in written directions using a RELATIVE `rframe`. The viewer attribute is used to capture the ID of the viewer entity[5].

(15) The [**school**$_{loc1}$] is to [**left of**$_{sig1}$] the [**post office**$_{loc2}$].
QSLINK(qs1, figureID=loc1, groundID=loc2, type=relational, relation=LEFT, rframe=relative, viewer='viewer')

Distance. A *distance link* relates two locations, spatial entities or states and specifies the distance between them. In addition to the IDs of the relevant spatial entities, this tag also includes `measure` and `unit` attributes.

Spatial Signal. Finally, we identify the signal (or trigger) that introduces a relation with the S_SIGNAL tag. A *spatial signal* is a relation word that locates a state or motion in a particular region or path, adds information to an existing event path, or provides information on the relation type of a spatial relation. We have seen examples of spatial signals in many of the previous examples in this section.

[5] While this ID is not part of the ISO-Space specification, we assume that we can inherit these ID references from an external named entity annotation. For the example, we assume that the sentence should be interpreted from the reader's perspective and we simply use 'viewer' to denote this.

4 Use and Applications

Earlier in this paper, we introduced four scenarios in which an ISO-Space annotation would be useful. We return to these below in order to demonstrate how the specification can deal with the specific problems of each scenario. The simplest case for ISO-Space is that of standard newswire because the specification was designed with this kind of text in mind. Recall the examples from above, repeated here with ISO-Space elements added:

> NYT: July 15, 2010
> [**CAIRO**$_{loc1}$] — A Libyan [**ship**$_{se1}$] that tried to [**break**$_{m1}$] Israels [**blockade**$_{se2}$] of [**Gaza**$_{loc2}$] [**docked**$_{m2}$] in the [**Egyptian port of El Arish**$_{loc3}$] on Thursday afternoon as the ship's sponsor [...] said that the [**boat**$_{se3}$] had [**shifted course**$_{m3}$] because the Israeli government agreed to allow Libya to support [**building**$_{e2}$] and [**reconstruction**$_{e3}$] in [**Gaza**$_{loc4}$].

Note that *ship* is annotated but not *Libyan* since this particular property would not help in any spatial inference. While the word *docked* is ambiguous between a state and a motion reading, in this context, it is clear that it behaves as a motion. Interestingly, *break* is annotated as a motion, but there are other complex issues at play here. This is a case where the interplay between ISO-Space and other annotation schemes is vital. ISO-Space can capture the meaning of "breaking a blockade", but subordination of this action into the modal context introduced by *tried* would come from the expressiveness of ISO-TimeML.

We next return to the scenario of text feeds from multiple viewers. ISO-Space provides for annotation of spatial relationships according to different frames of reference. Therefore, any text feeds from a GPS-identified viewer (author of the text) that contain point of view and relative frame of reference information can be appropriately annotation. If there are accompanying images or other sources of data, then the annotation data can be integrated into this additional data. Similar remarks hold for the related example in Figure 1.

Now consider the application of ISO-Space in the image annotation scenario in Figure 2. Once objects of interest in the image have been annotated, qualitative spatial links can be added to describe the relationships between these objects. In the example in Figure 2, three of the tagged objects in Figure 2 are buildings that can easily be geolocated. If our goal is to normalize the image data and textual descriptions provided from an annotator's relative point of view to an absolute frame of reference, the qualitative spatial links (QSLinks) contain the necessary information to do so. Therefore, if we can geolocate the viewer while also geolocating objects annotated from the image, the obvious absolute links can be generated.

More interestingly, however, given this grounding and the relational links provided by a corresponding text, we can now narrow the absolute position of objects that were not geolocated. For example, given the link below that is associated with a textual description from the annotation shown in example

(16), we can now determine the absolute position of the car (a non-geolocated object) more precisely.

(16) I see the Horticultural Hall behind the red car.
QSLink(qs1, figureID='Horticultural Hall', groundID='red car', type=relational, rframe=relative, viewer='viewer', type=BEHIND)

The remaining scenario involves directions for moving towards a specific goal. The language of such directions is actually quite different from the language used in the other three scenarios and, thus, represents a challenge for ISO-Space. In the next section, we discuss what the challenges are with this genre and present a possible treatment using ISO-Space.

5 Outstanding Issues and Concluding Remarks

We have seen in previous sections that while ISO-Space is intended to be rich enough for facilitating spatial awareness, it also allows extensions to its data categories by permitting elements from external ontologies to be plugged in, assuming these elements have a clear semantics. In this section, we introduce a number of issues that have not yet been adequately vetted by the developers of ISO-Space with the larger GIS and QSR communities. We begin by presenting a somewhat speculative extension for a specific genre: written directions to a specific location[6].

(17) To Airlie Center from Washington/National Airport
Take I-66 West to Exit 43A (Gainesville/Warrenton) and proceed South on Rt. 29 for approximately 10 miles. Take a right at the traffic light onto Colonial Road (Route 605). Colonial Road turns into Airlie Road, continue straight ahead on Route 605. Cross a one-lane stone bridge and take an immediate left into Airlie main entrance.

Given the elements introduced above in Section 3, it is possible to annotate verbs like *take* and *continue* in (17) as motions. But these are motions of a very particular kind because, unlike with most motions, a sequence of motions in directions is typically continuous. That is, the next motion starts off where the previous motion ended. In directions, a continuous composite directed path is built from segments of paths and maneuvers. A MANEUVER, adopted from [17], is a motion that is always associated with a TRAVERSAL element in the same way that a MOTION is associated with an EVENT_PATH. Traversals are composite parts of the entire route that is described. It is an abstract element that does not directly correspond with an extent from the text. Each maneuver in a route description introduces a traversal of a path and the traversal is defined by the path and the traversal's begin and end points on that path. The path is often,

[6] This extension was developed in the context of the Airlie ISO-Space Workshop held in 2010 (see www.iso-space.org for more information).

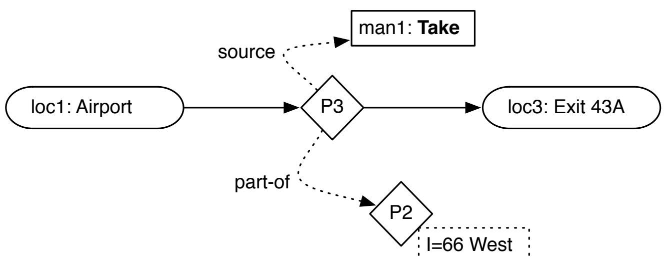
but not always, an explicit PATH element. Traversals are inherently directional and, like event paths, introduce a temporal component. In the simple example below, taken from (17), the maneuver *take* is associated with the traversal of a segment of *I-66*.

(18) [**Take**$_{ma1}$] [**I-66**$_{p2}$] West to [**Exit 43A**$_{loc3}$].

Maneuver(ma1, traversal=tr1)
Traversal(tr1, source=ma1, path=p2 begin=loc1, end=loc3)

Similar fragments can be created for the other sentences of (17) and all the fragments can be unified together into one large graph. The fragment of that graph corresponding to (18) is shown in (19).

(19)



The TRAVERSAL element has a similar ontological status as the EVENT_PATH element and it is debatable whether both elements are required. We introduced the element for its close association with the MANEUVER element. Not surprisingly, the set of attributes of TRAVERSAL is taken from the set defined on EVENT_PATH and includes attributes to store the identifiers of source, object, begin point and endpoint. Note that the object will point at the abstract entity that follows the direction. Given this discussion, there are reasons to think that maneuvers are a useful and expressive element for ISO-Space, but, at this point, it is not clear how they should be integrated into the rest of the specification.

Finally, other issues that remain unaddressed in the current version of the specification include the following[7]:

(20) a. Do we need a distinction between LOCATION and PLACE?
b. Should the granularity of the spatial relation be specified and, if so, how?
c. What spatial attributes associated with objects should be represented (e.g., shape, size)?
d. Should goals and intensional goals be distinguished?
 a. John left New York for Boston.
 b. The plane is on its way to Paris.
e. Should there be a distinction between PATH and PATH-SEGMENTS, when referring to parts of a route, for example?

[7] Initial discussion of these issues was begun at the Airlie Workshop.

f. What constitutes the core inventory of spatial relation values?

Many of these issues will most likely be resolved through the development of annotated corpus fragments, wherein alternative specification proposals can be tested for expressiveness, coverage, and ease of annotation. We are currently in the process of preparing such a corpus, while also developing the first fragment of a semantics for the abstract syntax presented here.

It should be noted that the selected corpora are all from written discourse. Clearly it would be beneficial to also consider other application areas including human-human or human-machine dialogues, for example car navigation systems, route planning, map searching, and navigating robots. In addition, all current corpora used are in English. In a later stage we will also include corpora from other languages since the ultimate idea is to have a multilingual annotation specification.

Acknowledgments

We would like to thank the members of the ISO-Space Working Group for their significant input to the current specification as well as the participants of the Airlie Workshop. We would also like to thank Anna Rumshisky for contributing to the work on the image annotation specification. Part of this research was funded under the NURI grant HM1582-08-1-0018 by the National Geospatial Agency.

References

1. Asher, N., Sablayrolles, P.: A typology and discourse for motion verbs and spatial pps in french. Journal of Semantics 12, 163–209 (1995)
2. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The berkeley framennet project. In: Proceedings of the COLING_ ACL. Montreal, Canada (1998)
3. Bateman, J.A., Hois, J., Ross, R., Tenbrink, T.: A linguistic ontology of space for natural language processing. Artificial Intelligence (2010)
4. Bennett, B., Galton, A.: A unifying semantics for time and events. Arificial Intelligence 153, 13–48 (2004)
5. Boas, H.C.: Frame semantics as a framework for describing polysemy and syntactic structures of english and german motion verbs in contrastive computational lexicography. In: Rayson, P., Wilson, A., McEnery, T., Hardie, A., Khoja, S. (eds.) Proceedings of the corpus linguistics 2001 conference. vol. 13. University Center for Computer corpus research on language, Lancaster, UK (2001)
6. Cappelle, B., Declerck, R.: Spatial and temporal boundedness in english motion events. Journal of Pragmatics 37(6), 889–917 (June 2005)
7. Egenhofer, M., Clementini, E., Di Felice, P.: Topological relations between regions with holes. International Journal of Geographical Information Systems 8(2), 129–142 (1994)
8. Fellbaum, C.: WordNet: An electronic lexical database. The MIT press (1998)
9. Frank, A.: Spatial ontology: A geographical information point of view. In: Stock, O. (ed.) Spatial and Temporal Reasoning, pp. 135–153. Springer Netherlands (1997)

10. Galton, A.: Towards a qualitative theory of movement. In: Frank, A.U., Kuhn, W. (eds.) Spatial Information Theory: A Theoretical Basis for GIS (Proceedings of International Conference COSIT'95). pp. 377–396. Springer-Verlag, Semmering, Austria (September 1995)
11. Galton, A.: Space, time and movement. In: Stock, O. (ed.) Spatial and Temporal Reasoning, pp. 321–352. Kluwer, Dordrecht (1997)
12. Grenon, P., Smith, B.: Snap and span: Towards dynamic spatial ontology. Spatial Cognition and Computation 4(1), 69–104 (2004)
13. Herring, J., Mark, D.M., (eds.), J.H., Mark, D.M., Egenhofer, M.J., Egenhofer, M.J.: The 9-intersection: Formalism and its use for natural-language spatial predicates (1994)
14. Herskovits, A.: Language and Spatial Cognition: an Interdisci- plinary Study of the Prepositions in English. Cambridge University Press (1986)
15. Jackendoff, R.: Semantics and Cognition. MIT Press (1983)
16. Kipper, K., Korhonen, A., Ryant, N., Palmer, M.: Extensive classifications of english verbs. In: Proceedings of the 12th EURALEX International Congress. Turin, Italy (2006)
17. Klippel, A., Hansen, S., Davies, J., Winter, S.: A high-level cognitive framework for route directions. In: Proceedings of the SSC 2005 Spatial Intelligence, Innovation and Praxis: The National Biennial Conference of the Spatial Science Institute. Melbourne (2005)
18. Kurata, Y., Egenhofer, M.: The 9+ intersection for topological relations between a directed line segment and a region. In: Gottfried, B. (ed.) Workshop on Behaviour and Monitoring Interpretation. pp. 62–76. Germany (September 2007)
19. Mani, I., Hitzeman, J., Richer, J., Harris, D., Quimby, R., Wellner, B.: Spatialml: Annotation scheme, corpora, and tools. In: Chair), N.C.C., Choukri, K., Maegaard, B., Mariani, J., Odjik, J., Piperidis, S., Tapias, D. (eds.) Proceedings of the Sixth International Language Resources and Evaluation (LREC'08). European Language Resources Association (ELRA), Marrakech, Morocco (may 2008), http://www.lrec-conf.org/proceedings/lrec2008/
20. Miller, G.: Wordnet: a lexical database for english. Communications of the ACM 38(11), 39–41 (1995)
21. MITRE: Spatialml: Annotation scheme for marking spatial expressions in natural language. http://sourceforge.net/projects/spatialml/ (December 2007)
22. Muller, P.: A qualitative theory of motion based on spatio-temporal primitives. In: Cohn, A.G., Schubert, L., Shapiro, S.C. (eds.) KR'98: Principles of Knowledge Representation and Reasoning, pp. 131–141. Morgan Kaufmann, San Francisco, California (1998)
23. Pustejovsky, J., Knippen, R., Littman, J., Saurí, R.: Temporal and event information in natural language text. Language Resources and Evaluation 39, 123–164 (May 2005)
24. Pustejovsky, J., Lee, K., Bunt, H., Romary, L.: Iso-timeml: An international standard for semantic annotation. In: LREC (2010)
25. Pustejovsky, J., Moszkowicz, J.L.: Integrating motion predicate classes with spatial and temporal annotations. In: Proceedings of COLING 2008. Manchester, UK (2008)
26. Pustejovsky, J., Moszkowicz, J.L., Verhagen, M.: Iso-space: The annotation of spatial information in language. In: Proceedings of ISA-6: ACL-ISO International Workshop on Semantic Annotation. Oxford, England (January 2011)

27. Randell, D., Cui, Z., Cohn, A.: A spatial logic based on regions and connections. In: Kaufmann, M. (ed.) Proceedings of the 3rd Internation Conference on Knowledge Representation and REasoning. pp. 165–176. San Mateo (1992)
28. Talmy, L.: How language structures space. In: Pick, H., Acredolo, L. (eds.) Spatial Orientation: Theory, Research, and Application. Plenum Press (1983)
29. Talmy, L.: Lexicalization patterns: Semantic structure in lexical forms. In: Shopen, T. (ed.) Language typology and semantic description Volume 3: Grammatical categories and the lexicon, pp. 36–149. Cambridge University Press (1985)
30. Vieu, L.: Sémantique des relations spatiales et inférences spatio-temporelles: une contribution à l'étude des structures formelles de l'espace en langage naturel. Ph.D. thesis, Université Paul Sabatier, Toulouse (1991)
31. Zaenen, A., Bobrow, D., Condoravdi, C.: The encoding of lexical implications in verbnet: Predications of change of location. In: Proceedings of LREC. Marakesh, Morocco (2008)