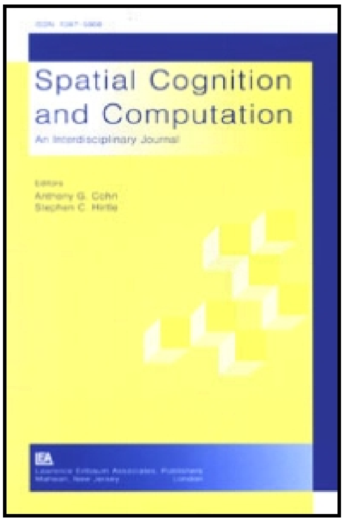

Spatial Cognition & Computation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653698

The Qualitative Spatial Dynamics of Motion in Language

James Pustejovsky[a]; Jessica L. Moszkowicz[a]
[a] Laboratory for Linguistics and Computation, Brandeis University, Waltham, Massachusetts, USA

To cite this Article Pustejovsky, James and Moszkowicz, Jessica L.(2011) 'The Qualitative Spatial Dynamics of Motion in Language', Spatial Cognition & Computation, 11: 1, 15 — 44
To link to this Article: DOI: 10.1080/13875868.2010.543497
URL: http://dx.doi.org/10.1080/13875868.2010.543497

Spatial Cognition & Computation, 11:15–44, 2011
Copyright © Taylor & Francis Group, LLC
ISSN: 1387-5868 print/1542-7633 online
DOI: 10.1080/13875868.2010.543497

The Qualitative Spatial Dynamics of Motion in Language

James Pustejovsky and Jessica L. Moszkowicz
Laboratory for Linguistics and Computation, Brandeis University,
Waltham, Massachusetts, USA

Abstract: In this paper, we discuss the strategies that languages employ to express motion, focusing on the distinction between path predicates, such as *enter, arrive*, and *leave* and manner-of-motion predicates, such as *walk*, *bike*, and *roll*. We present an overview of some qualitative spatiotemporal models of movement, and discuss their adequacy for capturing motion constructions in natural languages. Building on many aspects of these qualitative models, we introduce a framework within dynamic logic for the characterization of spatial change. This model, called Dynamic Interval Temporal Logic (DITL), is developed to analyze both classes of motion predicates, as well as complex compositional constructions involving spatial and manner Prepositional Phrases. Further, DITL serves as a semantics for a linguistically expressive markup language for annotating spatiotemporal information in text, called Spatiotemporal Markup Language (STML). We outline the syntax of this language, and discuss how DITL provides for a natural interpretation of the annotation specification for use in a variety of applications.

Keywords: spatial language, qualitative reasoning, common sense models of space, dynamic logic, temporal logic

1. INTRODUCTION

The interpretation of motion in language presents an interesting challenge to the qualitative spatial reasoning (QSR) community. Motion in natural language is generally viewed as encoding two aspects of meaning: *where* the movement is happening and *how* it is happening. Languages employ two very different strategies to accomplish this, a fact that has often been overlooked by the QSR community. Unfortunately, linguistic theories of motion have, likewise, largely overlooked many of the contributions from QSR. Namely, qualitative approaches not only help to ground linguistic expressions but also

Correspondence concerning this article should be addressed to Dr. James Pustejovsky, Department of Computer Science, MS-018, Brandeis University, 415 South Street, Waltham, MA 02454. E-mail: pustejovsky@gmail.com

embed them within richer and more expressive representational and reasoning frameworks.

In this paper, we formalize these two linguistic strategies for expressing motion in terms of distinct qualitative spatial representations. We exploit the fact that path constructions encode the *where* component of motion meaning by analyzing them as denoting movement relative to a *distinguished location*, a point or region on a path that is traversed by the moving object. Manner-of-motion verbs, as we shall see, do not make any explicit mention of a distinguished location, but still assume a change of location. This linguistic distinction has not been exploited for qualitative reasoning purposes. Galton, however, does make a crucial distinction between path movement and process movement, which is what we develop here as well (Galton, 1995). Motivated by a broader set of linguistic data and phenomena, we extend this to all aspects of language that compositionally make this same distinction, whether as a predicate, a Prepositional Phrase (PP), or other adjunct type. We will demonstrate how this distinction helps capture the compositional processes involved in combining the different aspects of meaning in motion expressions, an essential property of natural languages (Partee, 1984; Szabó, 2000; Werning, 2004).

Accounting for the language of motion is a vital component to the overall goal of creating a distinct, well-defined, and empirically-motivated layer of semantics for spatial language. Such a model has direct applications in many currently relevant tasks including attempts to support communication via natural language with human users of Geographic Information Systems (GIS). GIS traditionally includes fields such as cartography and surveying, but recently work in this area has expanded to include photogrammetry, remote sensing, spatial databases, spatial cognition, and spatial statistics. There is also interest in studying the underlying principles of geospatial technologies, including analysis and simulation of pedestrian movement. In fact, computing dynamic information has recently become an important part of emerging GIS technologies.

A further area in which a better understanding of the connection between natural language and formal representations of space is required 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. 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 such as the analysis of corresponding image data. 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 the present discussion, we focus specifically on the language of motion, a central issue to the more general problem of spatial language understanding.

To illustrate these problems, consider the following excerpt from a travel blog about biking through Central America, focusing on the distinction between path and manner-of-motion verbs, denoted below as p and m respectively:[1]

> John left$_p$ San Cristobal de Las Casas four days ago. He arrived$_p$ in Ocosingo that day. The next day, John biked$_m$ to Agua Azul and played in the waterfalls there for 4 hours. He spent the next day at the ruins of Palenque and drove$_m$ to the border with Guatemala the following day. (1)

In order to annotate the spatiotemporal elements of this text, we must be able to identify several kinds of information, including temporal and spatial expressions, as well as the predicate-argument structure. There are several existing annotation schemes for capturing these elements, including SpatialML (Mani et al., 2010) for locations and TimeML (Pustejovsky et al., 2003) for temporal information. A resource such as PropBank[2] can be used for understanding predicate-argument structure. Yet, none of these existing resources say much, if anything, about motion. To help remedy this, we develop a theory of motion based on qualitative spatial dynamics, which addresses the fundamental distinction in the way languages express motion, namely that between path and manner-of-motion constructions.

Path predicates introduce reference to a distinguished location such as in the sentence *He arrived in Ocosingo that day.* Pure manner-of-motion predicates do not make use of a distinguished location, as in *John biked all day*; they can, however, be used in a distinguished location interpretation by embedding the motion verb within a path construction, as seen in *John biked to Agua Azul.* We represent motion with Dynamic Interval Temporal Logic (DITL), which will serve as the semantics for STML, a spatial markup language that incorporates the annotations provided by SpatialML and TimeML, along with the ability to capture information about paths and the locations of events.

2. QUALITATIVE MODELS FOR MOTION IN LANGUAGE

In this section, we first illustrate the two major strategies that are employed in natural language to describe the movement of an object through space. We

---

[1] http://www.rideforclimate.com

[2] (Palmer, Gildea, & Kingsbury, 2003)

then explore how results from qualitative spatial reasoning can contribute to modeling the basic computational semantics of motion in language. For reasons of focus and space, we will ignore issues of movement with orientation or frame-of-reference information.[3] While obviously important for modeling linguistic expressions of motion, these issues should be addressed after the foundation has been laid for the basic semantics of movement.

Natural languages have two distinct strategies for expressing concepts of motion (Talmy, 1985): *path constructions* and *manner-of-motion constructions*.[4] The latter strategy can be seen in sentences such as those below in (2) (where m indicates a manner verb, and p indicates a path).

a. John hopped$_m$ [out of the room]$_p$.
b. Mary crawled$_m$ [to the window]$_p$. (2)

The path construction is illustrated with the following examples in (3):

a. John arrived$_p$ [by foot]$_m$.
b. John left$_p$ the building [running]$_m$. (3)

We can split languages broadly into the two classes. 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*. Languages that employ this construction include English, German, Russian, Swedish, Chinese, and others. 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*. Languages that employ this construction include Modern Greek, Spanish, Italian, Japanese, Turkish, Hindi, and others.

As observed in (2) and (3), English allows both constructions, and these are common in everyday language, such as the travel blog example in (1). For example, *bike* is a manner verb, but when used with a path PP-construction the sentence indicates both direction and path information. The verbs *arrive* and *leave* are both path verbs and give no information regarding the manner-of-motion, but manner adjunct PP-constructions provide further context.

Given these observations, in order to model motion in qualitative terms, we must track the property of an object's location as it changes over time. One way to represent object location is with a version of the Region Connection Calculus, such as RCC8 (Randell, Cui, & Cohn, 1992), which consists of eight jointly exhaustive and pairwise disjoint relations: disconnected (DC), externally connected (EC), partial overlap (PO), equal (EQ), tangential proper

---

[3] For a discussion of these issues, see (Freksa & Zimmermann, 1992; Noyon, Claramunt, & Devogele, 2007) and (Freksa, 1992; Mitra, 2004; Renz & Mitra, 2004).

[4] Subsequent work on this includes (Jackendoff, 1983; Talmy, 2000; Choi & Bowerman, 1991).

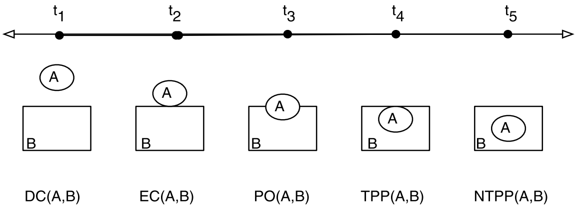

Figure 1. Galton analysis of *enter* using RCC8 relations.

part (TPP) and its inverse (TPPi), and non-tangential proper part (NTPP) and its inverse (NTPPi).

These relations provide the foundation for expressing simple topological relations between objects. To track changes in these relations, reference to some sort of temporal logic is needed. Galton discusses such a theory of change for motion using RCC relations (Galton, 1993, 1997), and develops these ideas more fully in (Galton, 2000). Recent work by Bhatt and Loke (2008) provides a somewhat related approach to modeling spatial change, where change of location (using RCC8 relations) is analyzed in the framework of the situation calculus. However, because the aim of that work is to address classic ramification and frame problems from Artificial Intelligence, their focus is not on capturing the semantics of motion in natural language.

Muller (1998) develops a theory of motion based on spatiotemporal primitives, using the mereotopology developed in (Asher & Vieu, 1995). This system is similar to RCC8 but adds the concept of open and closed sets (absent from RCC) as well as a set of temporal relations that include a relation of temporal connection, as well as the standard ordering relations. One aspect of this approach that is significant to our current discussion is that it clusters motion verbs in natural languages into distinct qualitative spatiotemporal representations.[5]

Galton also makes specific mention of how some natural language predicates can be interpreted within a qualitative model. He develops an analysis that embeds RCC8 relations within a temporal framework, where spatial relations are associated with a temporal index. The result is a logically grounded qualitative model of movement, as illustrated by the sequence of relations for the path predicate *enter* in Figure 1.

Work within the 9-Intersection calculus (9IC) (Egenhofer & Franzosa, 1991) has also been adopted to correlate with explicit spatial expressions in language, particularly the different ways lines intersecting with regions can be expressed (Egenhofer & Mark, 1995). The 9-Intersection Model for line-region relations is based on the intersections of the interiors, boundaries, and

---

[5] This classification is modified and extended somewhat in (Pustejovsky & Moszkowicz, 2008), where semantic considerations from (Asher & Sablayrolles, 1995) are incorporated into Muller's set.

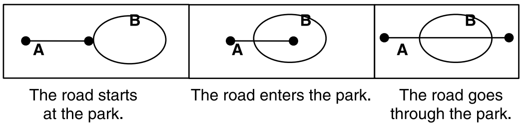

Figure 2. Possible linguistic correlates of some 9IC relations.

exteriors of a line (represented by A) and a region (represented by B), both interpreted as point sets, in the following matrix, where A^o represents the line interior, ∂A represents the line boundary, and A^- represents the line exterior, while B^o represents the region interior, ∂B represents the region boundary, and B^- represents the region exterior:

$$I(A,B) = \begin{pmatrix} A^o \cap B^o & A^o \cap \partial B & A^o \cap B^- \\ \partial A \cap B^o & \partial A \cap \partial B & \partial A \cap B^- \\ A^- \cap B^o & A^- \cap \partial B & A^- \cap B^- \end{pmatrix} \tag{4}$$

For example, imagine that A represents an abstraction of a road, while B represents a park, as in Figure 2.

Mark and Egenhofer (1995) report that specific line-region intersection values correspond to identifiable linguistic expressions denoting spatial configurations of lines to regions, as elicited by human subjects when shown configurations of lines and regions. Many of these expressions actually involve motion verbs, but are used to express static spatial relations, with what are called *fictive motion* constructions (Talmy, 2000). Unlike the RCC-based models mentioned above, however, there is no temporal information inherent in the representation of the spatial configurations between regions. Furthermore, as is clear from Figure 2, without direction, line-region intersection values cannot distinguish between "entering", "exiting", and so forth.

To solve problems related to this issue, Kurata and Egenhofer (2007) extend the 9IC, to the $9I^+$ calculus, where the notion of a *directed line* is introduced. Using this model, we can view a line, L, as having two distinct endpoints, $\partial_L L$ (left boundary) and $\partial_R L$ (right boundary). When intersected with a region, R, the resulting matrix, I^e, can be defined as the intersection between R and the two-boundaried line L shown below[6]:

$$I^e(L,R) = \begin{pmatrix} L^o \cap R^o & L^o \cap \partial R & L^o \cap R^- \\ \partial_L L \cap R^o & \partial_L L \cap \partial R & \partial_L L \cap R^- \\ \partial_R L \cap R^o & \partial_R L \cap \partial R & \partial_R L \cap R^- \\ L^- \cap R^o & L^- \cap \partial R & L^- \cap R^- \end{pmatrix} \tag{5}$$

This representation allows for a formal distinction between a "line pointing out of a region", and a "line pointing towards a region". However, in order

---

[6] The matrix used in our discussion is notationally different than what they present in their paper, for purposes of presentation.

to model change of location over time, the $9I^+$ matrix representations would have to be interpreted over temporal indexes.

One way to capture change of location using the $9I^+$ model might be to view a matrix as encoding the value of intersective relations from multiple temporal indexes. Motion would be read off the matrix as a temporal trace of the directed line-region intersection cell values, thereby allowing for interpretations of *leave* and *arrive*, for example.

But there is a problem with interpreting a directed LR-intersection matrix in this respect.[7] On this view, the verb *arrive*, for example, would correspond to $[\partial_L L \cap \partial R = 0]@t_1$, $[L^o \cap \partial R = 0]@t_2$, and $[\partial_R L \cap \partial R = 1]@t_3$. Assuming that the other entries in the relation matrix can assume any allowed value, then this description is underspecified, in that the motion could start in the interior or exterior of the region and end on the region boundary (similar remarks hold for *leave*). We can, however, solve this problem by using a point-line model, interpreted over explicit temporal indexes.

Consider the matrix below, showing the intersection of a point P and a line L, where the point P is indexed according to temporal indexes, t_1, t_2, and t_3, and the line L has the directed topological transformations, $\partial_L L$, the line's left boundary, L^o, the line's interior, and $\partial_R L$, the line's right boundary (we ignore L^- for the present discussion).

$$I(P,R) = \begin{pmatrix} P_1 \cap \partial_L L & P_1 \cap L^o & P_1 \cap \partial_R L \\ P_2 \cap \partial_L L & P_2 \cap L^o & P_2 \cap \partial_R L \\ P_3 \cap \partial_L L & P_3 \cap L^o & P_3 \cap \partial_R L \end{pmatrix} \tag{6}$$

Thus, when viewed as a Point-Line intersection over time, path predicates can be expressed in a snapshot model (Grenon & Smith, 2004), as with the verb *arrive*, shown in (7).

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7}$$

Hence, basic path verbs do seem to have a model in an extension of 9IC that incorporates explicit temporal indexing, using intersection relations with directed lines and points.

From this brief review, we see that relational models of spatial change can be fairly easily embedded within a temporal logic in order to account for basic linguistic expressions denoting change (e.g., (Galton, 2000; Bhatt & Loke, 2008; Muller, 1998)). Intersective models, on the other hand, must make explicit reference to temporal frames (indexes) as part of the intersecting values. The advantage of the directed LR model discussed above, is that there

---

[7] We would like to thank one of our reviewers for pointing out the inconsistencies with interpreting the directed LR intersection over temporal indexes.

is a reified spatial object that corresponds to the path, along which the the object (point) is moving, which is not the case in relational models, where no path region is reified. However, as Weghe, Kuijpers, and Bogaert (2005) point out, there is a further limitation in both basic RCC and intersection models, in that *disconnected from* (DC) relations are not differentiated, making it impossible to represent many concepts relating to movement towards or away from, as well as relative movement between, two objects. The Qualitative Trajectory Calculus (Weghe, 2004) overcomes this shortcoming by making comparisons between the positions of two objects at different moments in time. Partly based on the Double-Cross Calculus (Freksa & Zimmermann, 1992), it allows for the qualitative representation of varying values in DC relations between two objects (e.g., two objects approaching each other, or one object pulling further away from another, etc.). This is an expressive model and merits integration into a computational semantics for language, but this is a topic for future investigation.

Because the works reviewed here are primarily concerned with aspects of formal representation and reasoning over spatial calculi and not the linguistic expressions that denote such representations, it is not surprising that less attention has been paid to the compositional properties of how motion expressions are constructed in language. In the next section, we turn finally to this issue. We build on many of the ideas reviewed in this section and attempt to model the compositional aspects of motion in language, paying particular attention to the semantic distinction between manner-of-motion predicates and path predicates, as well as how they combine in language.

3. DYNAMIC INTERVAL TEMPORAL LOGIC

To adequately model the motion of objects as expressed in language, the representational framework should have at least two properties: (i) it should be inherently temporal; and (ii) it should accommodate change in the assignment of values to the relevant attributes being tracked, e.g., the location of an object.

One model that satisfies both of these properties for modeling motion is the situation calculus, as developed recently in (Bhatt & Loke, 2008). Situation calculus approaches to modeling change, by virtue of the temporal logic they assume, equate the meaning of the expression with its truth conditions, interpreted over the appropriate temporal frame. For example, a process and its effects are modeled as an axiom in the calculus, the instantiation of which is interpreted in the model over temporal indexes, which are inherent in the model. For this reason, temporal logics are often called *endogenous logics* (Pnueli, 1977).[8]

[8] Work by (Nr, Doherty, Gustafsson, Karlsson, & Kvarnstrom, 1998) and references therein attempt to represent pre- and post-conditions in change, within action logics and other models adopting Sandewall's features and fluents.

In natural languages, notions of time and temporality are encoded both implicitly in the tense and aspect system of the language (Comrie, 1985), (Mani, Pustejovsky, & Gaizauskas, 2005), as well as explicitly referenced through temporal prepositions and referring expressions (Pratt-Hartmann, 2005). Hence, there are both conceptual and linguistic motivations for reifying temporal indexes as first-class objects in the logic, as is done in the situation calculus and natural language event calculi (cf. (Bennett & Galton, 2000; Parsons, 1990; Pustejovsky, 1995).

It is also the case, however, that the notion of updating (changing) values associated with particular attributes of individuals is an inherent part of language. That is, many predicates in natural language reference an explicit change in the value of an object's attribute (e.g., *The temperature increased*, *The vase broke*, *John entered the room*). For this reason, dynamic logic has recently been applied to many aspects of linguistic reasoning and computation involving epistemic updates in dynamic contexts (cf. (Goldblatt, 1992; Harel, Kozen, & Tiuyn, 2000), and (Groenendijk & Stokhof, 1990)). It is this aspect of dynamic logic that is attractive for modeling linguistic constructions denoting change of state; namely, the property of update (e.g., change-of-location) is explicitly encoded in the logic. The discrete (step-by-step) simulation of change and iterated change of location developed below relates directly to (Grenon & Smith, 2004) and their temporalized construction of "snapshots".

In the remainder of this section, we outline a first-order fragment of a dynamic logic for encoding spatial change that we call *Dynamic Interval Temporal Logic (DITL)*, which combines both those aspects from temporal logic updating temporal information with change-of-state updates from dynamic logic. As such, this model meets the requirements outlined above. Then we demonstrate how this logic expresses both atomic motion and complex motion expressions in natural language, through complex predicative constructions as well as adjunct Prepositional Phrase constructions.

Within dynamic approaches to modeling updates, there is a distinction made between formulae, ϕ, and programs, π. A formula is interpreted as a classical propositional expression, with assignment of a truth value in a specific state in the model. For our purposes, a state is a set of propositions with assignments to variables at a specific time index. We can think of atomic programs as input/output relations, i.e., relations from states to states, and hence interpreted over an input/output state-state pairing. We will model "assignment-of-location" as an atomic first-order program, and, since the semantics of an atomic program is its input/output relations, we can treat change-of-location and other complex motion expressions as *compound programs*. The relation denoted by a compound program will be determined by the relations denoted by its atomic parts. This property, known as compositionality, makes dynamic logic attractive for modeling many natural language interpretations.

Recall the distinction between path and manner constructions observed above. Predicates making direct reference to a path, such as *arrive* or *leave*, specify a *distinguished location* along that path, either explicitly, as in *He arrived in Ocosingo that day*, or implicitly, as in *John left this morning*. Manner-of-motion predicates by themselves make no reference to any specific locations at all, as seen in *John biked all day*; they can, however, be used in a distinguished location interpretation by embedding the motion verb within a path construction, as seen in *John biked to Agua Azul.*

We can now develop these basic observations about motion predicates in dynamic terms. As mentioned above, there are two sets of symbols associated with dynamic logic, where S is the set of states: formulae ($[\![\phi]\!] \subseteq S$), and programs ($[\![\pi]\!] \subseteq S\ X\ S$).[9] For the present discussion, we limit our discussion of the formal mechanisms of the logic to those aspects relevant to modeling the two types of motion constructions introduced earlier in the paper. We assume the temporal operators normally associated with Linear Temporal Logic (LTL), such as Next ($\bigcirc$), All ($\Box$), Some ($\Diamond$), and Until ($\mathcal{U}$) (Pnueli, 1977; Vardi, 1996).[10] LTL is a discrete, linear model of time. This structure is represented by the model, $\mathcal{M} = \langle \mathbb{N}, I \rangle$, where $I : \mathbb{N} \mapsto 2^{\Sigma}$ maps each natural number (representing a moment in time) to a set of propositions, where Σ is the set of all atomic propositions.

First, we define the semantics of formulae in dynamic logic. Following standard assumptions within LTL, formulae have the following interpretations:

a. $\langle \mathcal{M}, i \rangle \models \phi$ iff $\langle \mathcal{M}, i \rangle \models \phi$
"ϕ holds now."
b. $\langle \mathcal{M}, i \rangle \models \bigcirc\phi$ iff $\langle \mathcal{M}, i+1 \rangle \models \phi$
"ϕ holds at the next time."
c. $\langle \mathcal{M}, i \rangle \models \Diamond\phi$ iff $\exists j[i \leq j \wedge \langle \mathcal{M}, j \rangle \models \phi]$
"ϕ holds at some time in the future." (8)
d. $\langle \mathcal{M}, i \rangle \models \Box\phi$ iff $\forall j[i \leq j \rightarrow \langle \mathcal{M}, j \rangle \models \phi]$
"ϕ holds for every time in the future."
e. $\langle \mathcal{M}, i \rangle \models \phi\, \mathcal{U} \psi$ iff $\exists j[j \geq i \wedge \langle \mathcal{M}, j \rangle \models \psi \wedge \forall k[i \leq k < j \rightarrow$
$\langle \mathcal{M}, k \rangle \models \phi]$ "ϕ holds until ψ starts to hold."

Within dynamic logic, every program is interpreted with an input state s_1 and output state s_2. The program constructions that are most relevant to our discussion include: atomic programs, sequences of programs, testing a formula, iteration, and reporting the output of a program. These constructions along with their corresponding interpretations in LTL are given below, where interpretations in the model are evaluated relative to pairs of temporal indexes,

---

[9] We assume the syntax of Propositional Dynamic Logic (PDL) (Harel et al., 2000).

[10] Cf. also (Kröger & Merz, 2008; Allen, 1984; Moszkowski, 1986; Manna & Pnueli, 1995). We will avoid the use of temporal operators in the following discussion when not necessary.

(i, j). Note that the letters a and b are used to represent atomic programs while α and β represent compound programs.

a. Any atomic program, a, is a program;
“Execute program a”.
$\langle \mathcal{M}, (i, i+1)\rangle \models a$ iff $\langle \mathcal{M}, i\rangle \models s_1 \wedge \langle \mathcal{M}, i+1\rangle \models s_2$
b. If a and b are atomic programs, then $a; b$ is a compound program called a *sequence*;
“Execute a, then execute b”;
$\langle \mathcal{M}, (i, j)\rangle \models a; b$ iff $\exists k[[i \leq k \leq j \wedge \langle \mathcal{M}, (i, k)\rangle$
$\models a \wedge \langle \mathcal{M}, (k, j)\rangle \models b]$;
i.e. $k = i + 1$ and $j = i + 2$.
c. If α and β are programs, then $\alpha; \beta$ is a program called a *sequence*;
“Execute α, then execute β”;
$\langle \mathcal{M}, (i, j)\rangle \models \alpha; \beta$ iff $\exists k[[i \leq k \leq j \wedge \langle \mathcal{M}, (i, k)\rangle$
$\models \alpha \wedge \langle \mathcal{M}, (k, j)\rangle \models \beta]$
d. If ϕ is a formula, then $\phi?$ is a program called a *test*;
“Check the truth value of ϕ, and proceed if ϕ is true, fail if false[11]”;
$\langle \mathcal{M}, (i, i+1)\rangle \models s_1 \rightarrow \top$
e. If a is a program, then a^* is a program called *Kleene iteration*;
“Execute a zero or more times.”
$\langle \mathcal{M}, (i, j)\rangle \models a^*$ iff $\forall k[i \leq k \leq j \rightarrow \langle \mathcal{M}, (k, k+1)\rangle \models a]$
f. If a is an atomic program and ϕ is a formula, then $[a]\phi$ is a formula;
“It is always the case that after executing a, ϕ is true.”
$\langle \mathcal{M}, (i, i+1)\rangle \models [a]\phi$ iff $\langle \mathcal{M}, i\rangle \models \bigcirc\phi$
g. If α is a program and ϕ is a formula, then $[\alpha]\phi$ is a formula;
“It is always the case that after executing α, ϕ is true.”
$\langle \mathcal{M}, (i, j)\rangle \models [\alpha]\phi$ iff $\langle \mathcal{M}, j-1\rangle \models \bigcirc\phi$

(9)

To illustrate better how dynamic logic expressions are interpreted in a linear temporal logic, consider the compound program, $a^2; b; c$, as executed in the diagram in Figure 3. From (9g), we see that ϕ is a formula that holds at time j. Since we are associating “one step of a program, π_i” directly with one movement of the time index, we can gloss the formula $[\alpha]\phi$ as defined in Figure 3 as follows, along with other equivalences:

a. $[\alpha]\phi$ means “Every execution of $a^2; b; c$ results in ϕ”.
b. $[c]\phi$ is equivalent to $\bigcirc\phi$ at time $j - 1$.
c. $\langle \pi_i^* \rangle\phi$ is equivalent to $\Diamond\phi$ at time i, where π_i is any atomic program.[12]

(10)

---

[11] This will have the effect of a skip operation to the next program in the sequence.

[12] As in modal logic, the “diamond” operator is the dual of “box”, where $\langle\alpha\rangle\phi$ means, “There is a computation of α that terminates in a state satisfying ϕ.”

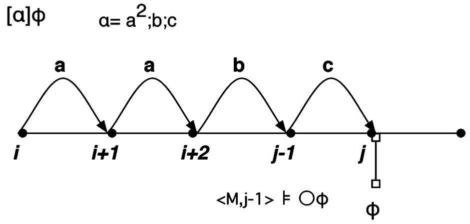

Figure 3. Tracing a compound program.

In order to capture the change in an attribute that an object can undergo in a dynamic context, we must obviously enrich the logic presented above to a first-order language. First-order models require the addition of assignment functions associated with each state at a given time, in order to keep track of the values bound to variables in the expressions being interpreted (e.g., $x \mapsto george$, $y \mapsto boston$, $z \mapsto loc_3$).

For the present discussion, we assume the following atomic program, *variable assignment*, which associates a specific value to a variable. This requires that we extend the model to pairs of assignment functions (or valuations) (u, v), in addition to temporal index pairs, (i, j). That is, every program, a, in our language, $a \in \pi$, is evaluated with respect to a pair of states, $[\![\pi]\!] \subseteq S\ X\ S$, and with each state there is an assignment function. Hence, in order to evaluate a program, a pair of assignment functions is required.

$$\begin{array}{l} \text{If } x \text{ and } y \text{ are variables, then } x := y \text{ is an atomic program.} \\ \quad \text{“}x \text{ assumes the value given to } y \text{ in the next state.”} \\ \quad \langle \mathcal{M}, (i, i+1), (u, u[x/u(y)]) \rangle \models x := y \\ \quad \text{iff } \langle \mathcal{M}, i, u \rangle \models s_1 \wedge \langle \mathcal{M}, i+1, u[x/u(y)] \rangle \models x = y \end{array} \tag{11}$$

Example (11) states that the value of the variable x is newly assigned as y, as interpreted over a pair of model assignment functions, u, the input state assignment, and $u[x/u(y)]$, the output state assignment, which is exactly like u except that the value it assigns to x has been replaced with y.[13] For example, assigning the location of an object x as l_1, is written as the atomic program, $loc(x) := l_1$.

Using the tools developed above, let us return to our concerns about the semantics of motion predicates in natural language. The most significant observation from our previous discussion is that path verbs such as *arrive* and *leave* are inherently different from basic manner-of-motion predicates, such as *move*, *roll*, and *walk*, in that they make explicit reference to the location that is being moved away from or toward along an explicit path. Manner verbs,

[13] See (Groenendijk & Stokhof, 1989) and (Eijck & Stokhof, 2005) for discussion of dynamic assignment strategies in computational semantics.

as we shall see, still assume a change of location while making no explicit mention of a distinguished location. Within the model being developed here, this distinction is operationally very clear:

a. PATH VERBS involve movement relative to a distinguished location; hence, they involve a program *testing* for that location of the moving object;
b. MANNER-OF-MOTION VERBS involve no distinguished locations; they involve *assignments* of locations of the moving object from state to state.

(12)

We now fully develop how DITL accounts for each of these constructions.

3.1. Semantics of Manner-of-Motion Predicates

The most basic program of motion, a "change-of-location", involves a variable assignment and reassignment to the value of an identified spatial attribute: e.g., $loc(x) := y$.[14] This requires reference to not only a pair of temporal indexes (i, j) along with an intermediate index, k, that pairs with both of them, (i, k) and (k, j), but also reference to a pair of assignment functions (u, v) and an intermediate assignment, w, that pairs with each of them, (u, w) and (w, v). We define BASIC CHANGE OF LOCATION, *change_loc$_{bas}$*, below.

$$\begin{array}{ll} \text{a. } change\_loc_{bas}(x) =_{df} & loc(x) := y \; ; y := z, y \neq z \\ & \langle \mathcal{M}, (i, j), (u, v) \rangle \models loc(x) := y \; ; y := z, \;\; y \neq z \text{ iff} \\ & \exists k \exists w [[i \leq k \leq j \wedge (u, w) \wedge (w, v) \wedge \langle \mathcal{M}, (i, k), (u, w) \rangle \\ & \models loc(x) := y \wedge \langle \mathcal{M}, (k, j), (w, v) \rangle \models y := z, \;\; y \neq z] \end{array} \quad (13)$$

With the definition of basic change of location given in (13), we can now define the general change-of-location predicate we will use in subsequent discussion, where there is an assignment of a location that is changed, and then Kleene iterated.[15]

$$change\_loc(x) =_{df} \; loc(x) := y \; ; (y := z, y \neq z)^{+} \quad (14)$$

For modeling motion predicates such as *walk*, *drive*, and other manner verbs, however, we need yet another constraint, in order to give direction or orientation to the movement. Here we make use of the distance constraint as employed in (Weghe et al., 2007), where we measure relative distance between

---

[14] We focus on the single spatial attribute of *location* in this paper. Conceptually, this treatment is close to Galton's (Galton, 2000) analysis of movement as change of position and to (Bhatt & Loke, 2008) and their definition of primitive change of spatial relationship between two objects.

[15] We say Kleene iterated because α^{+} indicates one application of α followed by $\alpha*$.

distinct assigned values to the location of an object. Let $d(l_1|t_1, l_2|t_2)$ denote the Cartesian distance between two temporal indexed points. If we identify the starting location of any directed motion as a point, b, then we can ensure motion away from that point using the linear distance constraint in (15).

$$d(b|t_i, y|t_i) \ < \ d(b|t_{i+1}, z|t_{i+1}) \tag{15}$$

With this defined, we arrive at the necessary constraints for directed motion within a dynamic framework, illustrated below:

DIRECTED MOTION:

a. Assign a value, y, to the location of the moving object, x.
$loc(x) := y$
b. Name this value b (this will be the beginning of the movement);
$b := y$
c. Then, reassign the value of y to z, whose distance from b has increased, $d(b, y) < d(b, z)$;
$y := z,\ d(b|t_i, y|t_i) \ < \ d(b|t_{i+1}, z|t_{i+1})$
d. Kleene iterate step (c).

(16)

This is rendered as the DITL program in (17).

$$move_{dir}(x) =_{df} \ loc(x) := y, b := y \ ; (y := z, y \neq z, d(b, y) < d(b, z))^+ \tag{17}$$

To illustrate this, consider the meaning of the manner-of-motion verb *roll*, as used in (18).

The ball *rolled* quickly along the street. (18)

Ignoring for now the semantic contribution made by the specific manner of the movement (i.e., "rolling" versus "sliding"), the verb *roll* denotes a directed motion verb. Let us consider the valuation of this predicate that brings the ball to a specific location, l_3, as visualized in the diagram in Figure 4.

We assume the initial location of the ball, x is assigned as l_1. We designate this initial location as the begin point, b. Then we change the location of the ball by reassigning the value of $loc(x)$. At each iteration of the process, we check that the distance constraint is satisifed, namely that the distance from b to the newly assigned location, l_k, is growing. At time $j - 1$, the reassignment of the location, $loc(x) := z$ is evaluated relative to the temporal index pair $(j - 1, j)$ and the assignment function pair $(v - 1, v)$, in our model, $\mathcal{M}$, returning $loc(x) = l_3$ at time j.

The definition of directed motion given in (17) will work for linear movement, but as pointed out in (Weghe et al., 2007), this will not work for directed motion involving 2D movement given the definition of the distance constraint. For example, it will be unable to account for the initially increasing

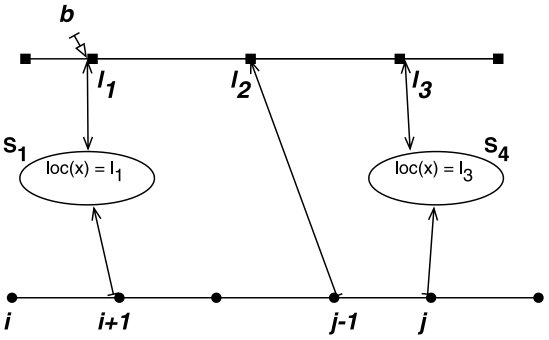

Figure 4. Directed motion.

and subsequent decreasing relative distance as an object proceeds around the boundary of a region (19a), or for an object in a circular motion (19b).

a. John walked the perimeter of the building.
b. Mary walked around the lake. (19)

In both these cases, distance must be measured along the structure of a path, p, and not simply relative to the begin point, b, of the movement. In these examples, the spatial configuration of the path is determined by the meaning of the direct object Noun Phrase in (19a) and the spatial Prepositional Phrase in (19b).

Accounting for directed motion in 2D space in complex configurations (such as circles and polygons) is beyond the scope of the present discussion. However, this raises the issue that all manner-of-motion predicates leave a trail of the motion along an implicit path, as measured over time. We will refer to this as *motion leaving a trail*, and define it operationally below:

MOTION LEAVING A TRAIL:
a. Assign a value, y, to the location of the moving object, x.
$$loc(x) := y$$
b. Name this value b (this will be the beginning of the movement);
$$b := y$$
c. Initiate a path p that is a list, starting at b;
$$p := (b)$$
d. Then, reassign the value of y to z, where $y \neq z$
$$y := z, y \neq z$$
e. Add the reassigned value of y to path p;
$$p := (p, z)$$
e. Kleene iterate steps (d) and (e); (20)

A manner verb, as shown above, does not presuppose a path along which the motion is traversed. Rather, the motion creates the path incrementally and

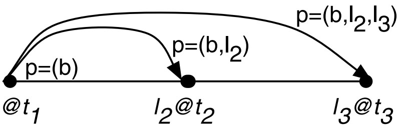

Figure 5. Directed motion leaving a trail.

dynamically. The above operational constraints are captured by the following DITL expression, called $move_{tr}$:[16]

$$\begin{aligned} move_{tr}(x) =_{df} \; & loc(x) := y, b := y, p := (b); \\ & (y := z, y \neq z, p := (p, z))^{+} \end{aligned} \tag{21}$$

Now we can combine directed motion and motion leaving a trail, to give us a directed motion with a trail.

$$\begin{aligned} move_{dir+tr}(x) =_{df} \; & loc(x) := y, b := y, p := (b) \; ; \\ & (y := z, y \neq z, p := (p, z), d(b, y) < d(b, z))^{+} \end{aligned} \tag{22}$$

To illustrate this motion type, notice how the path in Figure 5 is iteratively expanded in the following trace of $move_{dir+tr}$.

By reifying the path created by the motion, we are now able to quantify over it, as illustrated in the examples below:

(23)
a. The ball rolled 20 feet.
$\exists p \exists x \exists e[[roll(e, x, p) \wedge ball(x) \wedge length(p) = [20, foot]]$
b. John biked for 5 miles.
$\exists p \exists e[[bike(e, j, p) \wedge length(p) = [5, mile]]$

In sum, we have shown how manner-of-motion predicates always consist of an initial motion followed by zero or more iterations of that same motion. As a result of this movement, a path is created, tracing the steps of the object in motion. Further, we defined directed motion with the help of a simple distance constraint.

3.2. Semantics of Path Predicates

While all motion involves a change of location, path verbs denote movement relative to a distinguished location, a point or region on a path that is traversed by the moving object. The change of location that is denoted by a path predicate is evaluated relative to the distinguished location along the

---

[16] Notice that this definition allows some fairly diverse movement types (such as oscillations and rotations), since it only requires a Markov change in location; that is, location values can be revisited arbitrarily.

designated path. For example, the manner verb *walk* as in *Mary walked yesterday* was analyzed above as an iterated directed motion, with no specific location referenced for the change of location. A path verb such as *enter*, however, as in *Mary entered the store* designates a distinguished region, *the store*, and evaluates motion on a path relative to that region.

In dynamic terms, what the verb *enter* is doing is designating a location, and then conditionalizing any directed motion towards that location. In other words, a path verb incorporates a program that *tests* whether the current location of the moving object is the same as this distinguished location on this path. If it is not, then movement is made towards that location.

Recall the definition of *test* presented above:

a. If ϕ is a formula, then $\phi?$ is a program called a *test*;
"Test ϕ, and proceed if ϕ is true, fail if false"; (24)

A first-order test involves checking the value of the variable associated with an object attribute, such as $loc(x)$. For example, consider the verb *arrive* as used in *John arrived in Boston*. Given the goal location that is mentioned in the sentence (i.e., *Boston*), the appropriate test in this case would be that in (25).

a. $(loc(j) \neq boston)?$
"Is it not the case that John's location is Boston?" (25)

If this test succeeds, then we want something (α) to happen that changes the value of this attribute, until its negation succeeds, i.e., (26):

a. $(loc(j) = boston)?$
"Is it the case that John's location is Boston?" (26)

The α, of course, is a movement predicate, as defined earlier in this section (e.g., *change_loc* or $move_{dir}$). Putting these components together, we have an operational definition for path predicates such as *enter* and *arrive*:

PATH PREDICATE:

a. Identify a distinguished location (or region), d, on a path, p, denoted by the interval $[p_1, p_2]$. Assume d is either the begin point or end point of p;
$p := [d, p_2]$ or $p := [p_1, d]$
b. Test the location of the moving object, x against the distinguished location, d;
$(loc(x) \neq d)?$
b. If (b) is true, execute some movement, α;
c. Kleene iterate steps (a) and (b);
d. Test the negation of the formula in (a);
$(loc(x) = d)?$

(27)

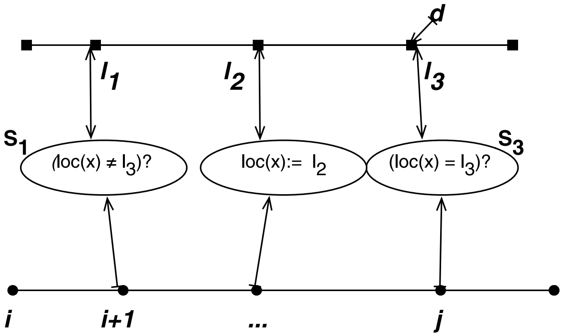

Figure 6. Path verb interpretation.

Note that the above definition works for testing the location of an object when the distinguished location is the goal, as with *enter*. When the distinguished location references the source of the movement, as with the verb *exit*, the test will have to be appropriately defined. Given this observation, the definition above translates to the two DITL expressions below, where for ease of exposition, we distinguish *arriving*-path predicates ($move_{a\_path}$), such as *arrive* or *enter*, from *leaving*-path predicates ($move_{l\_path}$), such as *depart*, *exit*, and *leave*.

$$\begin{aligned}&\text{a. } move_{a\_path}(x,d) =_{df} \; p := [p_1, d]; (loc(x) \neq d?; move_{dir}(x))^*;\\ &\quad loc(x) = d?\\ &\text{b. } move_{l\_path}(x,d) =_{df} \; p := [d, p_2]; (loc(x) = d?; move_{dir}(x))^*;\\ &\quad loc(x) \neq d?\end{aligned} \tag{28}$$

Figure 6 illustrates a trace of the semantics of a path predicate program corresponding to the verb *arrive*. The initial component of the program tests the location of the object relative to the distinguished location, d, which is l_3. After the initial test, reassignment of the location of the object x is performed, iteratively, until the test against the distinguished location is satisfied, at time j.

In sum, in this section we have shown how path predicates in language involve a distinguished region or location on a designated path. Any change in location of an object is made relative to this distinguished location by virtue of testing that object's location against this value.

3.3. Compositional Constructions

In the previous discussion, we defined the semantics for basic change-of-location (*change_loc*), and used this to define both directed movement ($move_{dir}$) and path predicates ($move_{a\_path}$ and $move_{l\_path}$). These two strategies for denoting motion can be combined, so that both kinds of information can be encoded in the same sentence. There are two possible compositional constructions for combining path and manner information. As we saw in the previous section, English allows both of these constructions, though some languages prefer or prohibit one or the other.

a. Use a *manner-of-motion* verb (*run*, *bike*) with a *path* adjunct (Prepositional Phrase indicating spatial path information);
b. Use a *path* verb (*enter*, *arrive*) with a *manner* adjunct (Prepositional Phrase indicating the manner of the movement). (29)

Consider first strategy (29a). Given a manner verb construction as used in (30),

John biked$_m$ in the morning. (30)

we can modify the manner process with a spatial Prepositional Phrase, *to Agua Azul*, denoting the end point of the motion, as in (31).

John biked$_m$ in the morning [to Agua Azul]$_p$. (31)

As mentioned above, the PP *to Agua Azul* introduces both an *explicit path*, p, and the distinguished location of this path (d), namely, Agua Azul, a.

To account for this construction compositionally, we need to analyze the path-inducing preposition, *to*, as a relation between locations and programs that move the object to that location (cf. (Pustejovsky, 1991a, 1995) for an event semantic treatment of this view). This is illustrated below in (32).

$$to(\pi(x), d) =_{df} \quad p := [p_1, d]; (loc(x) \neq d?; \pi(x))^*; loc(x) = d? \qquad (32)$$

This states that a path Prepositional Phrase, such as *to Agua Azul*, introduces a path variable, p, along with a distinguished location, d (which is the object of the preposition itself), and establishes a testing environment, within which a directed movement predicate, $\pi(x)$ is placed; in other words, this embeds the location assignment semantics from *bike* within the testing environment created by *to Agua Azul*.

Figure 7 demonstrates how manner verbs are embedded within a path construction created by a spatial PP. In this figure, the initial and final test conditions ($loc(x) \neq d?$ and $loc(x) = d?$) refer to the tests on the location of John relative to Agua Azul, viz. $loc(j) \neq a?$ and $loc(j) = a?$, respectively. The intermediate program, π, in this case, denotes the directed manner-of-motion predicate, *bike*.

The DITL expression associated with this sentence is given below.

$$p := [y, d], loc(j) := y, d := a; (loc(j) \neq a?; bike(j))^*; loc(j) = a? \qquad (33)$$

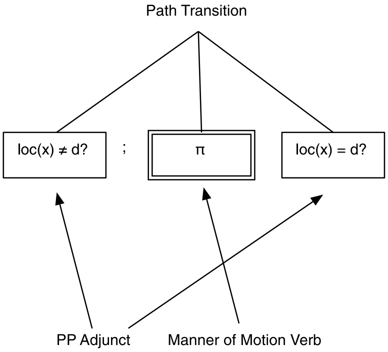

Figure 7. Manner Verb + Path PP: *bike to Agua Azul.*

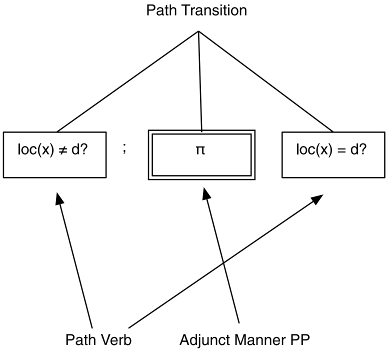

Figure 8. Path Verb + Manner PP: *leave by foot.*

This first defines a path with Agua Azul, a, assigned as the distinguished location. It then checks John's location against a, and executes iterations of $bike(j)$ until the location test is satisfied. This construction will explain the semantics of all such sentences involving path phrases added to manner verbs, as shown below.

a. John walked$_m$ [to the ruins]$_p$.
b. The baby crawled$_m$ [to the window]$_p$. (34)

Now let us consider strategy (29b), where a path predicate such as *leave* can incorporate manner information, thereby indicating both the path traversed as well as the manner of the movement. Consider a path verb construction as used in (35) below.

John left$_p$ Ocasingo this afternoon. (35)

Manner can be incorporated as the means by which the path is traversed with the use of a manner adjunct Prepositional Phrase, as illustrated in (36).

John left$_p$ Ocosingo this afternoon [by foot]$_m$. (36)

The resulting composition is illustrated in Figure 8.

It should be pointed out that some prepositions such as *from* always introduce an assignment at the start of the interpretation of a motion. For example, *John walked* is a simple manner-of-motion predicate, but adding *from* as in *John walked from the store* introduces an initial assignment. Such initial assignment prepositions have the interpretation given in (37), where a distinguished location, d, is identified by assignment rather than testing, and then acts as a function over a motion predicate, $\pi(x)$.

$$from(\pi(x), d) =_{df} \ loc(x) := d; (\pi(x))^* \tag{37}$$

Hence, the sentence in (38) has a distinguished beginning for the motion, but it does not denote a classic path construction, since there is no test of the distinguished location, as defined above.

John walked$_m$ [from the store] for thirty minutes. (38)

In the next section, we show how DITL influences the design of a spatiotemporal markup language for motion, STML, which is important for the development of spatial processing algorithms.

4. SEMANTIC ANNOTATION OF MOTION IN TEXT

The development of DITL provides us with an expressive language that addresses many of the problems associated with motion as introduced earlier in this paper. We now return to the travel blog text to see how these different aspects of motion can be associated with a markup language, thereby providing a vocabulary for DITL. Such an annotation language, which we call Spatiotemporal Markup Language (STML) allows us to identify where the different components of motion are expressed linguistically in text.

Good annotations are expressive reflections of the semantic content of a particular aspect of the text. Typically XML-based, they map easily to logical representations as well as to interoperable interchange formats that function as standards. This allows for their utility as data structures for mapping, visualization, and other grounding applications. Annotation of text is done in the service of developing algorithms (rule-based and machine learning) for automatic extraction of such information. But developing an annotation language or scheme is the first step, one that reflects the distribution of information in the language.

There are several elements of the text that will serve as ingredients for a DITL representation. Consider the following sentence from the travel blog:

The next day, John biked to Agua Azul and played in the waterfalls there for 4 hours. (39)

Our goal is to identify what spatiotemporal information is needed in order to track the entity in motion, *John*. First, there are explicit mentions of locations such as *Agua Azul* and *the waterfalls* that need to be indentified. As we saw in the previous section, such locations obviously play an important role in DITL representations as either location assignments or tests. For annotating locations such as these, STML builds on the SpatialML specification.

The focus of SpatialML is to identify spatial locations mentioned in text 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 has attributes `type` (country, continent, populated place, building, etc.), `country`, `gazref` (a reference to a gazetteer

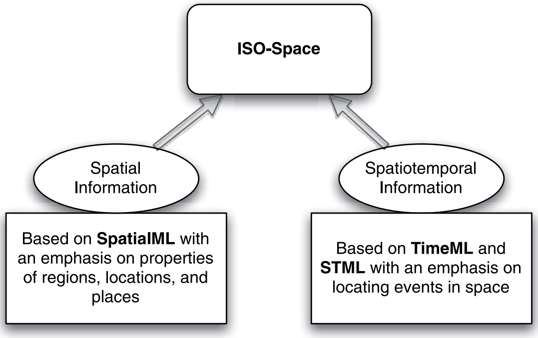

Figure 9. ISO-space components.

entry) and `latlong` (latitude and longitude values). 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. The link types for the LINK tag are adopted from the RCC8 version of the Region Connection Calculus. The SpatialML link types are mostly topological in nature and include: IN (tangential and non-tangential proper parts), EC (extended connection), DC (discrete connection), PO (partial overlap), EQ (equality), and NR (near), the only non-topological type.

SpatialML is one of the cornerstones of ISO-Space (Moszkowicz & Pustejovsky, 2010), a new standard being developed within the ISO TC 37/SC 4 for spatial and spatiotemporal annotation.[17] ISO-Space incorporates and improves on the annotation that SpatialML provides by enriching its spatial expressiveness. Specifically, the annotation of locations in ISO-Space will encode spatial properties such as topological relations between objects, orientation and metric relations between objects, the shape of an object, the size of an object, elevation, geopolitical entities, granularity, and aggregates and distributed objects.[18] The details of this aspect of ISO-Space are beyond the scope of this paper, but ISO-Space also incorporates the annotation provided by STML in order to capture spatiotemporal information. Figure 9 shows how the different specifications (SpatialML, STML, and TimeML) are incorporated in ISO-Space. For the remainder of this section, we turn our attention to how STML captures spatiotemporal information.

As we have seen, the recognition of spatial entities is an important component of understanding a text (Mani, Hitzeman, & Clark, 2008), but simply identifying fixed geospatial regions and specific "facilities" is not enough to achieve a complete representation of all the spatial phenomena present, since it leaves out one of the most crucial aspects of spatial information, motion.

---

[17] J. Pustejovsky is editor of the Work Item within ISO for this effort.

[18] ISO-Space also draws heavily on (Bateman, Hois, Ross, & Tenbrink, 2010) for spatially relevant categories.

To capture motion, we must integrate temporal and spatial information with the lexical semantics of motion predicates and prepositions.

Any annotation scheme designed to capture spatiotemporal information must have a temporal component. Indeed, the logical forms associated with a DITL representation are inherently temporal in nature. STML annotation makes use of TimeML for capturing temporal referencing and ordering.[19] The vocabulary of STML, however, does not itself include temporal primitives since these can be inherited from TimeML.

TimeML, which is now an ISO standard (ISO-TimeML), is a representation scheme for capturing the way temporal information is expressed in text. The basic elements of a TimeML annotation are temporal expressions such as dates, times, and durations, and events that can be anchored or ordered to those expressions or with respect to each other. Once these temporal objects are captured, they are related to each other by way of a temporal link. TimeML's temporal relations are based on Allen's 13 basic relations (Allen, 1984) and include before, simultaneous, includes, begins, ends, as well as their inverses and an identity relation.[20]

ISO-TimeML annotation identifies both motion predicates such as *biked* and non-motion predicates such as *played* as events. Examples such as *play*, while not involving any motion, do involve internal movement as defined in Muller's classification of motion (Muller, 1998). Such events are still important in the spatial understanding of a text, especially when they are explicitly related to a location. This is in fact the case for the *played* event which is situated at *the waterfalls*. STML includes a link called EVENT_LOCATION that relates this kind of event to a specific location.

Change of location predicates such as *arrive*, *leave*, and *bike*, are captured with the MOTION tag in STML, which is based on TimeML's EVENT tag. They also introduce an EVENT_PATH tag, which reflects the way motion is represented in DITL. That is, regardless of whether the predicate is a path or manner-of-motion construction, all motion predicates introduce a trail that is referred to as a path in DITL. In STML, a path is a special kind of region that can have a begin point and an end point, though these points can be underspecified. In fact, manner-of-motion predicates that appear without any path adjunct will have an EVENT_PATH with "unknown" as both the begin and end locations. Paths can be introduced explicitly in the text as in *John met Mary along the way*, but, in such a case, the text *the way* is captured in a LOCATION tag and, therefore, annotated in ISO-Space rather than STML. When a path is introduced by an explicit path predicate or referenced in a path Prepositional Phrase, as in *John biked to Agua Azul*, it is these constructions that are

---

[19] Capturing explicit temporal expressions such as *the next day* and *4 hours* allows us to ground the annotation on a timeline just as capturing explicit locations is needed to ground the annotation on a map.

[20] In addition to temporal links, ISO-TimeML includes subordinating links that are used to capture information about irrealis events. This allows temporal links to be created even when the participating events may or may not have happened.

captured in STML. Note that STML makes a distinction between manner-of-motion and paths, but does not know where these came from linguistically. DITL, however, has a compositional semantics that captures this information.

The role that spatial prepositions (*to*, *from*, *in*, etc.) play in STML is particularly motivated by DITL. Whether the preposition performs an initial assignment or introduces a test, all spatial prepositions have the effect of adding information to the EVENT_PATH associated with the motion predicate. Prepositions like these are captured in STML with the S_SIGNAL tag. The STML annotation of *John biked to Agua Azul* is given in example (40b). We also provide the DITL representation for this sentence in (40) for comparison.[21]

a. *John biked to Agua Azul.*
b. `<MOTION mid="m1" extent="biked" type="manner" / >`
`<S_SIGNAL ssid="ss1" extent="to" / >`
`<LOCATION lid="l1" extent="Agua Azul" / >`
`<EVENT_PATH epid="ep1" source="m1" start_locationID=`
`"unknown" end_locationID="l1" signalID="ss1" / >`
c. $p := [y, d], loc(j) := y, d := a; (loc(j) \neq a?; bike(j))^*;$
$loc(j) = a?$

(40)

When spatial prepositions such as *in* appear with path predicates, they also add information to the EVENT_PATH that is introduced by the predicate. In fact, the annotations of *John biked to Agua Azul* and *John arrived in Agua Azul* are very similar, as shown below in example (41). This follows directly from what DITL says about these constructions, namely, that path constructions and compositional constructions are both modeled in the same way.

a. *John arrived in Agua Azul.*
b. `<MOTION mid="m1" extent="arrived" type="path" / >`
`<S_SIGNAL ssid="ss1" extent="in" / >`
`<LOCATION lid="l1" extent="Agua Azul" / >`
`<EVENT_PATH epid="ep1" source="m1"` (41)
`start_locationID="unknown"`
`end_locationID="l1" signalID="ss1" / >`
c. $p := [y, d], loc(j) := y, d := a; (loc(j) \neq a?; move_{dir}(j))^*;$
$loc(j) = a?$

The specification of STML has been influenced by DITL in the following ways: (i) change of location predicates, whether they be path or manner-of-motion constructions, are captured with the MOTION tag, (ii) spatial prepositions that are used in motion constructions are captured with the S_SIGNAL

---

[21] Note that we do not include the entire annotation of locations or motion events in our examples here since this part of the annotation is handled by ISO-Space and TimeML, respectively.

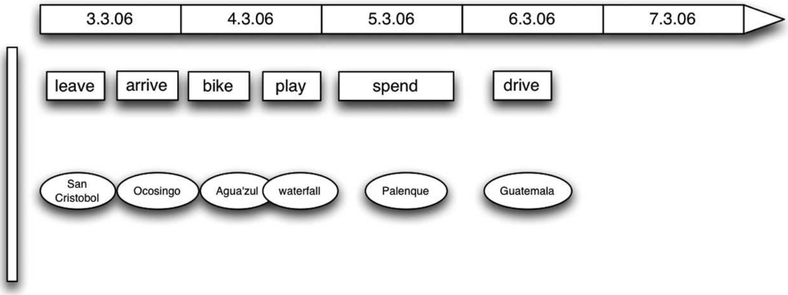

Figure 10. Ordering of events with corresponding locations.

tag,[22] and (iii) all motion events introduce an EVENT_PATH tag to the annotation, which describes the trace of the motion. Another benefit of annotating a text with STML is that the annotation can also be grounded in time and on a map. Recall that TimeML gives us the times, events, and ordering of those events, possibly anchored in time. STML goes a step further by also relating locations to those events as shown in Figure 10. In addition, STML, together with a complete ISO-Space annotation, also gives us path information and metric grounding. Figure 11 shows a schematic configuration of the metric constraints overlaid with an image of a map of the area. To actually ground the events on a map, ISO-Space uses the attributes associated with each location to connect to a resource such as the Keyhole Markup Language (KML), Google's file format for displaying geographic data in Google Earth or Google Maps. The details of this mapping, however, are still in development and beyond the scope of this paper.

Beyond the goal of automatically processing the spatiotemporal information in text, there are several additional issues that must still be addressed. For example, as the QSR community has observed, e.g., Weghe, Cohn, Bogaert, and Maeyer (2004); Hornsby and Cole (2007), motion constructions often involve more than a single entity in motion, such as a person *chasing* or *following* someone, or a car *overtaking* or *being passed by* another car. Since DITL and STML are both emerging resources, we have purposefully focused on simplistic examples for the present discussion. Extensions to this work that account for more complex constructions are currently being developed.

Another important area of research for spatial annotation is that of how scale and granularity issues are represented in STML and any subsequent grounded representation such as the one shown in Figure 11. For example, the travel blog excerpt tells us that John spent some time at the waterfalls that are located in Agua Azul. A complete ISO-Space annotation of the text will include tags that provide the relationship between the waterfalls and Agua

---

[22] Spatial prepositions that are used to describe locations such as *behind* in *behind the store* are captured in ISO-Space with the S_FUNCTION tag, but these are beyond the purview of STML.

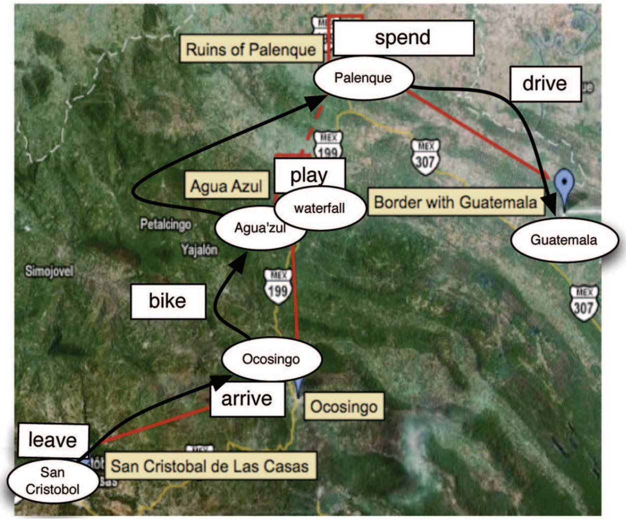

Figure 11. Schematic configuration of events and places with map. (Figure available in color online.)

Azul, but how to create an adequate representation of this change in scale in the grounded representation remains an open question. In fact, deciding on the most helpful ways to represent spatiotemporal information on a map, along with how to talk about this information verbally, is an important area of research that is still in its infancy.

5. CONCLUSION

In this paper we present a computational semantics for motion as expressed in natural languages, based in part on formal models used within the qualitative spatiotemporal reasoning community. We embedded the representation of spatial change of an object within a first-order modal dynamic logic, called Dynamic Interval Temporal Logic (DITL), and demonstrated how this language is able to naturally represent the two major strategies for encoding motion in language: path predicates and manner-of-motion predicates. This framework offers a compositional semantics that makes the distinction between motion constructions in language both operationally and denotationally transparent. It also explains the semantics of compositional constructions that combine manner-of-motion predicates with spatial prepositions such as *to* and *from*. The resulting interpretation, then, essentially mimics the behavior of path constructions. Currently, we are working to enrich the expressiveness of the language to account for orientation and frame-of-reference variables in motion descriptions.

DITL serves a dual purpose as it provides a new way to analyze motion as expressed in language, while also motivating how spatially relevant information in text should be annotated, in order to capture objects in motion. STML is a markup language that takes advantage of existing resources for annotating space and time, but also includes new elements for annotating motion as suggested by the representation presented here. The combination of a motion annotation with DITL as its semantics affords us an important tool in our understanding of the qualitative spatial dynamics of motion. Currently, the focus of STML is quite narrow since the development of the specification is housed in the more general ISO-Space project, which strives to capture all spatial information in text. STML is responsible for annotating spatiotemporal information such as the motion constructions that DITL represents and for anchoring non-motion events in space. Future work will include the development of spatial processing algorithms using this specification and DITL to automatically capture locations, paths, and motion constructions in text.

ACKNOWLEDGMENTS

We would like to thank Marc Verhagen and Anna Rumshisky for help in the preparation of this work, as well as Inderjeet Mani, Elisabetta Jezek, Annie Zaenen, and the participants of Dagstuhl Seminar 10131 for useful discussion. We would also like to express great appreciation for the thoughtful and constructive comments of the reviewers of this work. Portions of this work were presented at the 2009 Stanford Workshop on Spatial Relations, the 2009 AAAI Spring Symposium on Benchmarking Qualitative Spatiotemporal Systems, as well as the 2009 COSIT Conference in Aber W'rach, France. All remaining errors are, of course, the responsibility of the authors. This work was supported in part by a grant to James Pustejovsky by NGA HM1582-07-1-2037.

REFERENCES

Allen, J. (1984). Towards a general theory of action and time. *Artificial Intelligence, 23,* 123–154.

Asher, N., & Sablayrolles, P. (1995). A typology and discourse for motion verbs and spatial pps in French. *Journal of Semantics, 12,* 163–209.

Asher, N., & Vieu, L. (1995). Towards a geometry of common sense: a semantics and a complete axiomatisation of merotopology. In *Proceedings of ijcai95.* Montreal, Canada.

Bateman, J. A., Hois, J., Ross, R., & Tenbrink, T. (2010). A linguistic ontology of space for natural language processing. *Artificial Intelligence,* in press.

Bennett, B., & Galton, A. (2004). A unifying semantics for time and events. *Artificial Intelligence, 153,* 13–48.

Bhatt, M., & Loke, S. (2008). Modelling dynamic spatial systems in the situation calculus. *Spatial Cognition and Computation, 153,* 86–130.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in english and korean: The influence of language-specific lexicalization patterns. *Cognition, 41*(1–3), 83–121.

Comrie, B. (1985). *Tense.* Cambridge University Press.

Egenhofer, M., & Franzosa, R. (1991). Point-set topological spatial relations. *International Journal of Geographical Information Science, 5*(2), 161–174.

Egenhofer, M., & Mark, D. (1995). Modeling conceptual neighborhoods of topological line-region relations. *International Journal of Geographical Information Systems, 9*(5), 555–565.

Eijck, J. V., & Stokhof, M. (2006). The gamut of dynamic logics. In *Handbook of the History of Logic,* volume 6 (pp. 499–600). Elsevier.

Freksa, C. (1992). 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* (pp. 162–178). Pisa, Italy.

Freksa, C., & Zimmermann, K. (1992). On the utilization of spatial structures for cognitively plausible and efficient reasoning. In *Proc. of the conf. on systems, man, and cybernetics* (pp. 261–266). Chicago, USA.

Galton, A. (1993). Towards an integrated logic of space, time, and motion. In R. Bajcsy (Ed.), *Proceedings of the thirteenth international joint conference on artificial intelligence (ijcai'93)* (pp. 1550–1555). San Mateo: Morgan Kaufmann.

Galton, A. (1995, September). 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)* (pp. 377–396). Semmering, Austria: Springer-Verlag.

Galton, A. (1997). Space, time and movement. In O. Stock (Ed.), *Spatial and temporal reasoning* (pp. 321–352). Dordrecht: Kluwer.

Galton, A. (2000). Qualitative spatial change. Oxford, UK: Oxford University Press.

Goldblatt, R. (1992). Logics of time and computation (2nd ed.). CSLI Lecture Notes 7.

Grenon, P., & Smith, B. (2004). Snap and span: Towards dynamic spatial ontology. *Spatial Cognition and Computation, 4*(1), 69–104.

Groenendijk, J., & Stokhof, M. (1989). Type-shifting rules and the semantics of interrogatives. In B. P. Gennaro Chierchia & R. Turner (Eds.), *Properties, types and meaning* (Vol. 2, pp. 21–68). Dordrecht: Kluwer.

Groenendijk, J., & Stokhof, M. (1990). Dynamic predicate logic. *Linguistics and Philosophy, 14,* 39–100.

Harel, D., Kozen, D., & Tiuyn, J. (2000). *Dynamic logic* (1st ed.). Cambridge: The MIT Press.

Hornsby, K. S., & Cole, S. J. (2007). Modeling moving geospatial objects from an event-based perspective. *T. GIS, 11*(4), 555–573.

Jackendoff, R. (1983). Semantics and cognition. Cambridge: MIT Press.

Kröger, F., & Merz, S. (2008). *Temporal logic and state systems.* Springer-Verlag.

Kurata, Y., & Egenhofer, M. (2007, September). The 9+ intersection for topological relations between a directed line segment and a region. In B. Gottfried (Ed.), *Workshop on behaviour and monitoring interpretation* (pp. 62–76). Germany.

Mani, I., Doran, C., Harris, D., Hitzeman, J., Quimby, R., Richer, J., et al. (2010, September). Spatialml: annotation scheme, resources, and evaluation. *Language Resources and Evaluation, 44*(3), 263–280.

Mani, I., Hitzeman, J., & Clark, C. (2008). Annotating natural language geographic references. In *Workshop on methodologies and resources for processing spatial language.* Marrakesh, Morocco.

Mani, I., Pustejovsky, J., & Gaizauskas, R. (2005). *The language of time: A reader.* Oxford University Press.

Manna, Z., & Pnueli, A. (1995). *Temporal verification of reactive systems: Safty.* New York: Springer.

Mark, D., & Egenhofer, M. (1995). Topology of prototypical spatial relations between lines and regions in English and Spanish. In *Proceedings of the Twelfth International Symposium on Computer-Assisted Cartography,* volume 4 (pp. 245–254).

Mitra, D. (2004). Modeling and reasoning with star calculus: An extended abstract. *Eighth International Symposium on AI and Mathematics.* Fort Lauderdale, USA.

Moszkowicz, J. L., & Pustejovsky, J. (2010). Iso-space: Towards a spatial annotation framework for natural language. *Processing Romanian in Multilingual, Interoperational and Scalable Environments.*

Moszkowski, B. (1986). *Executing temporal logic programs.* Cambridge: Cambridge University Press.

Muller, P. (1998). 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* (pp. 131–141). San Francisco: Morgan Kaufmann.

Noyon, V., Claramunt, C., & Devogele, T. (2007). A relative representation of trajectories in geographical spaces. *GeoInformatica, 11*(4). Hingham: Kluwer Academic Publishers.

Nr, V., Doherty, P., Gustafsson, J., Karlsson, L., & Kvarnstrom, J. (1998). *Temporal action logics (tal): Language specification and tutorial.*

Palmer, M., Gildea, D., & Kingsbury, P. (2003). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics.*

Parsons, T. (1990). *Events in the semantics of English. a study in subatomic semantics.* Cambridge: MIT Press.

Partee, B. (1984). Compositionality. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics* (pp. 281–312). Dordrecht: Floris Publications.

Pnueli, A. (1977). The temporal logic of programs. In *18th Annual Symposium on Foundations on Computer Science.* Providence, USA.

Pratt-Hartmann, I. (2005). Temporal prepositions and their logic. *Artificial Intelligence, 166,* 1–36.

Pustejovsky, J. (1991a). The generative lexicon. *Computational Linguistics, 17*(4), 409–441.

Pustejovsky, J. (1991b). The syntax of event structure. *Cognition, 1*(41), 47–81.

Pustejovsky, J. (1995). *The generative lexicon.* Cambridge, MA: MIT Press.

Pustejovsky, J., Castano, J., Ingria, R., Saurí, R., Gaizauskas, R., Setzer, A., et al. (2003). Timeml: Robust specification of event and temporal expressions in text. In *Iwcs-5, fifth international workshop on computational semantics.*

Pustejovsky, J., & Moszkowicz, J. L. (2008). Integrating motion predicate classes with spatial and temporal annotations. In *Proceedings of Coling 2008,* Manchester, UK.

Randell, D., Cui, Z., & Cohn, A. (1992). A spatial logic based on regions and connections. In M. Kaufmann (Ed.), *Proceedings of the 3rd international conference on knowledge representation and reasoning* (pp. 165–176). San Mateo, USA.

Renz, J., & Mitra, D. (2004). Qualitative direction calculi with arbitrary granularity. In *Proceedings of the 8th Pacific rim international conference on artificial intelligence* (pp. 65–74). Springer.

Szabó, Z. G. (2000). *Problems of compositionality.* New York: Garland.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and semantic description volume 3: Grammatical categories and the lexicon* (pp. 36–149). Cambridge: Cambridge University Press.

Talmy, L. (2000). *Towards a cognitive semantics.* Cambridge: MIT Press.

Vardi, M. (1996). An automata-theoretic approach to linear temporal logic. *Logics for Concurrency. Structure versus Automata, LNCS, 1043,* 238–266.

Weghe, N. V. D. (2004). *Representing and reasoning about moving objects: A qualitative approach.* Phd thesis, Ghent University, Belgium.

Weghe, N. V. D., Bogaert, P., Cohn, A. G., Delafontaine, M., Temmerman, L. D., Neutens, T., et al. (2007). *How to handle incomplete knowledge concerning moving objects.* Germany: Osnabrüuck.

Weghe, N. V. D., Cohn, A., Bogaert, P., & Maeyer, P. D. (2004). Representation of moving objects along a road network. In *Proc. of geoinformatics* (pp. 187–197). Sweden: Gävle.

Weghe, N. V. D., Kuijpers, B., & Bogaert, P. (2005). A qualitative trajectory calculus and the composition of its relations. In *Proc. of geos* (pp. 60–76). Springer-Verlag.

Werning, M. (2004). Compositionality, context, categories and the indeterminacy of translation. *Erkenntnis, 2*(60), 145–178.