Integrating Motion Predicate Classes with Spatial and Temporal Annotations

James Pustejovsky
Computer Science Department
Brandeis University
jamesp@cs.brandeis.edu

Jessica L. Moszkowicz
Computer Science Department
Brandeis University
jlittman@cs.brandeis.edu

Abstract

We propose a spatio-temporal markup for the annotation of motion predicates in text, informed by a lexical semantic classification of these verbs. We incorporate this classification within a spatial event structure, based on Generative Lexicon Theory. We discuss how the spatial event structure suggests changes to annotation systems designed solely for temporal or spatial phenomena, resulting in spatio-temporal annotation.

1 Introduction and Motivation

The recognition of spatial entities in natural language is an important component of understanding a text (Mani et al., 2008). However, simply identifying fixed geospatial regions and specific "facilities" is not enough to achieve a complete representation of all the spatial phenomena present. In fact, this leaves out one of the most crucial aspects of spatial information, motion. To capture motion, we must integrate temporal and spatial information with the lexical semantics of motion predicates and prepositions.

The goal of this research is to further the representational support for spatio-temporal reasoning from natural language text in the service of practical applications. To create such support, we propose to use lexical resources for motion predicates to integrate two existing annotation schemes, SpatialML and TimeML, creating a representation that captures, in a fine-grained manner, the movement of individuals through spatial and temporal indexes. This work is part of a larger effort to automate such annotation and reasoning over natural language documents using symbolic and machine learning methods.

In this paper, we investigate different resources and annotations for spatio-temporal information. In section 2, we describe some of the resources we employed for our investigation. Section 3 elaborates on the classes we focus on as we work towards developing a classification for the purpose of annotating motion predicates, which we discuss in section 4.

2 Previous Work on Motion Classifications in Language

There has been considerable research on the linguistic behavior of spatial predicates and prepositions in language (e.g., (Jackendoff, 1983), (Herskovits, 1986), (Boas, 2001), (Cappelle and Declerck, 2005)). Within qualitative spatial reasoning (QSR), work has recently started to focus on incorporating mereo-topological concepts into the calculus of relations between regions. The most successful of these is the Regional Connection Calculus, or RCC (Randell et al., 1992). RCC8 and other systems like it do an adequate job of representing static information about space, but they cannot help us deal with motion, since that task requires a temporal component. Galton ((Galton, 1993; Galton, 1997)) began work on a commonsense theory of motion, but this work did not focus on merging temporal and spatial phenomena. Muller (Muller, 1998), however, proposes just such a system with his qualitative theory of motion based on spatio-temporal primitives. The result of Muller's system is a set of six motion classes: *leave*, *hit*, *reach*, *external*, *internal*, and *cross*.

Asher and Sablayrolles offer their own account

Coling 2008: Companion volume – Posters and Demonstrations, pages 95–98
Manchester, August 2008

of motion verbs and spatial prepositional phrases in French (Asher and Sablayrolles, 1995). They propose ten groups of motion verbs as follows: *s'approcher* (to approach), *arriver* (to arrive), *entrer* (to enter), *se poser* (to alight), *s'éloigner* (to distance oneself from), *partir* (to leave), *sortir* (to go out), *décoller* (to take off), *passer (par)* (to go through), *dévier* (to deviate). This verb classification is more fine-grained than Muller's.

While Muller, Asher, Sablayrolles, and Vieu among others have focused on the formal semantics of motion, other work has been done to represent motion in the FrameNet (Baker et al., 1998) and VerbNet (Kipper et al., 2006) projects. The Motion frame is a high level frame in the FrameNet hierarchy. It is defined as "Some entity (Theme) starts out in one place (Source) and ends up in some other place (Goal), having covered some space between the two (Path)."

To explore VerbNet's take on motion predicates, we mapped Asher and Sablayrolles' verbs to VerbNet classes. The mapping revealed that, while many of the motion predicates we care about have specific classes in VerbNet, it is not always clear what these classes have in common unless we look to FrameNet to find a higher level representation.

3 Classifying spatio-temporal predicates

Following (Muller, 1998), (Vieu, 1991), and (Asher and Sablayrolles, 1995), we assume spatial variables are incorporated into the representation of motion predicates in language. For this paper, we generally follow (Muller, 2002) by representing the individuals participating in spatial relations as spatio-temporal regions ($s{\cdot}i$). For modeling motion, however, we restrict our discussion to spatio-temporal regions occupied by physical matter denoted by the type $s \cdot i \cdot p$.

For this work, we performed several mappings between Muller, Asher and Sablayrolles, and FrameNet. The result of this mapping was a group of classes based largely on Muller's classifications with some very slight modifications detailed in the table below. The spatial event structures for each of these classes will describe their formal semantics (as in Figure 1 below).

In addition to these classes, we model the spatial semantics of prepositions, following (Asher and Sablayrolles, 1995), generally. Because of space limitations, we will not discuss the contritbution of prepositional semantics in this paper.

Move	run, fly, drive
Move_External	drive around, pass
Move_Internal	walk around the room
Leave	leave, desert
Reach	arrive, enter, reach
Detach	take off, disconnect, pull away
Hit	land, hit
Follow	follow, chase
Deviate	flee, run from
Stay	remain, stay

Table 1: Motion Classes

There is a complex interaction between a motion verb class and the interpretation of its arguments. For example, not all regions are occupied by the extent of physical matter (see above), but there are some objects which are properly both physical and spatial, such as the concept *building*. Notice the ambiguity inherent in the statement below, where both **Move_Internal** and **Move_External** are possible interpretations.

(1) The man walked around the building.

This is due to the semantic nature of *building* as both a physical object with extent, and also as a volume/aperture.

To model the mapping of objects to specific argument and event structures, we adopt the framework of Generative Lexicon Theory (GL). The notion of "polarity" in the (Muller, 1998) sense is quite similar to the semantic effect brought about by event headedness in (Pustejovsky, 1995). GL provides an explicitly typed argument structure, a typed subeventual structure, and a predicative body, which we will use to express RCC8 relations. For example, a representation of the Spatial Event Structure (SES) for the motion predicate *leave* is illustrated below in Figure 1. Note that the statement *Polarity=initial* is equivalent to saying *Head=left*. The relation BT is shorthand for boundary transition, which is composed of the following RCC8 relations: TPP, O, and EC.

Each motion class in Table 1 maps to a unique predicative body (qualia structure) in the spatial event structure for a verb. We demonstrate below how these representations are then embedded in the annotation of a text as RCC8 relations in a modified SpatialML/TimeML format, called Spatio-temporal Markup (STM).

The robustness of the mapping from the motion classes in Table 1 to FrameNet is currently being

$$
\begin{bmatrix}
\textbf{leave} & \\
\text{ARGSTR} = & \begin{bmatrix} \text{ARG1} = x\textbf{:}\ s \cdot i \cdot p \\ \text{ARG2} = y\textbf{:}\ s \cdot i \end{bmatrix} \\
\text{EVENTSTR} = & \begin{bmatrix} \text{E}_1 = \textbf{e}_1\textbf{:process} \\ \text{E}_2 = \textbf{e}_2\textbf{:state} \\ \text{RESTR} = <_{\propto} \\ \text{POLARITY} = \textbf{initial} \end{bmatrix} \\
\text{QUALIA} = & \begin{bmatrix} \text{AGENTIVE} = \textbf{NTTP(e}_1\textbf{,x,y)} \\ \wedge\ \textbf{BT(e}_1\textbf{,x.y)} \\ \text{FORMAL} = \textbf{DC(e}_2\textbf{,x,y)} \end{bmatrix}
\end{bmatrix}
$$

Figure 1: Spatial Event Structure

tested and evaluated.

4 Spatio-temporal Annotation

Throughout the development of the classification described here, we have tried to focus on how the classification will impact the task of annotating spatio-temporal information in text. There are currently two distinct annotation schemes for spatial and temporal information. If we are to successfully capture motion phenomena in text, these annotations must be merged just as a topological base and a temporal calculus need to be combined to model motion predicates.

TimeML (Pustejovsky et al., 2003) is an annotation scheme for representing temporal information 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.

For the annotation of spatial information, SpatialML (MITRE, 2007) is being developed. The focus of SpatialML is the markup of spatial locations that can be integrated with additional resources such as databases that provide information about a given domain (e.g. physical feature databases, gazetteers).

While SpatialML does a reasonable job of capturing locations in space, it cannot model moving objects such as people and cars, and in fact lacks a mechanism for capturing motion since predicates are not annotated. As we saw above, the motion event must be captured in addition to the locations involved in the motion. The development of our motion classification also reveals that the participants of the event are also needed, even if they are not spatial locations. For example, in *John flew to Boston*, *John* must be included in the motion annotation because he is a moving object.

We can enhance the spatio-temporal information from SpatialML and TimeML with lexical semantic information from a lexicon of motion predicates, resulting in annotation that is rich enough to be able to (1) infer motion of individuals invovled in specific events; and (2) to compose motions to create motion sequences (cf. (Pustejovsky and Verhagen, 2007)).

To create a spatio-temporal markup, TimeML and SpatialML must be enriched so that they can adequately capture the motion predicates we are discussing. TimeML will already annotate motion predicates as events, but to truly reveal the motion involved, additional attributes must be added to account for the beginning, middle, and end points of the event. The spatio-temporal annotation will require a new kind of spatial link to capture motion paths. Essentially, the motion path will combine the event information from TimeML with the spatial information from SpatialML. This motion path is at the core of a spatio-temporal markup or STM. The spatial event structure described in the previous section motivates the construction of the STM.

The concept of polarity or headedness will also motivate some aspects of the annotation. Depending on the polarity of the motion predicate or spatial prepositional phrase, the annotator will know to look for the source or goal of the event in the text and include that in the motion path.

The exact details of the resulting spatio-temporal markup are still under development. However, the following examples give an idea of how motion class information allow us to integrate spatial and temporal annotation.

(2) *John drove from Boston to NY on Thursday.*

```
<MOVER id=0>John</MOVER>
<EVENT id=1 tense=past start=t1 end=t2>drove</EVENT>
<SIGNAL id=2 type=spatial polarity=initial>from</SIGNAL>
<PLACE id=3>Boston</PLACE>
<SIGNAL id=4 type=spatial polarity=final>to</SIGNAL>
<PLACE id=5>New York</PLACE>
<SIGNAL id=6 type=temporal>on</SIGNAL>
<TIMEX3 id=7>Thursday</TIMEX3>
<TLINK eventID=1 timeID=7 relType=INCLUDES
   signalID=6 />
<MOTION eventID=1 moverID=0 source=3
   sourceTime=t1 sourceSignal=2 goal=5
   goalTime=t2 goalSignal=4 class=MOVE/>
```

(3) *John left Boston for New York.*

```
<MOVER id=0>John</MOVER>
<EVENT id=1 tense=past start=t1 end=t2
   polarity=intitial>left</EVENT>
<PLACE id=2>Boston</PLACE>
<SIGNAL id=3 type=spatial
   polarity=final>for</SIGNAL>
<PLACE id=4>New York</PLACE>
<MOTION eventID=1 moverID=0 source=2 sourceTime=t1
   goal=4 goalTime=t2 goalSignal=3 class=LEAVE/>
```

The `Motion` tag in the above examples tells us the class of the motion predicate. This provides a link to both the spatial event structure (as in Figure 1) and a spatio-temporal markup, which embeds the annotaton of a text as RCC8 relations in a modified SpatialML/TimeML format. The example in 4 shows the STM for *leave*:

$$(4)\quad \begin{bmatrix} \textbf{motion} \\ \text{TYPE} = \textbf{leave} \\ \text{EVENTID} = \mathbf{e} \\ \text{MOVERID} = \mathbf{x} \\ \text{SOURCE} = \mathbf{l}_1 \\ \text{SOURCETIME} = \mathbf{t}_1 \\ \text{GOAL} = \mathbf{l}_2 \\ \text{GOALTIME} = \mathbf{t}_2 \end{bmatrix} \Longrightarrow \begin{bmatrix} IN(t_1, x, l_1) \\ IN(t_2, x, l_2) \\ DC(t_2, x, l_1) \end{bmatrix}$$

This STM indicates what additional information is needed for the spatio-temporal annotation. In the case of example 3, three temporally anchored SpatialML `link` tags are indiated for each of the RCC8-like relations to the second part of the STM:

(5)
```
<LINK linkType=IN source=0 target=2 time=t1/>
<LINK linkType=DC source=0 target=2 time=t2/>
<LINK linkType=IN source=0 target=4 time=t2/>
```

These links that can be automatically generated at a later stage in the annotation set up the locations of moving objects at given times. The first link in example 5 reveals that the moving object *John* was in *Boston* at time *t1*, which is the start time of the motion given in the annotation.

5 Conclusion

In this paper, we investigate how an expressive classification for verbs of motion can be used to integrate spatial and temporal annotation information, in order to represent objects in motion, as expressed in text. We adopt a modified version of the classifications of verbs of motion in (Muller, 1998) and (Asher and Sablayrolles, 1995) and demonstrated how verb classes are mapped to RCC8+1 relations in a temporally anchored SpatialML. We are currently evaluating the reliability of the FrameNet encoding of motion predicates, and are developing algorithms for translating lexical structures to Spatio-temporal markup.

References

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

Baker, C., C. Fillmore, and J. Lowe. 1998. The berkeley framenet project. In *Proceedings of the COLING_ACL*, Montreal, Canada.

Boas, H. 2001. Frame semantics as a framework for describing polysemy and syntactic structures of english and german motion verbs in contrastive computational lexicography. Rayson, P., A. Wilson, T. McEnery, A. Hardie, and S. Khoja, eds., *Corpus linguistics 2001* , vol.13, Lancaster, UK.

Cappelle, B. and R. Declerck. 2005. Spatial and temporal boundedness in english motion events. *Journal of Pragmatics*, 37(6):889–917, June.

Galton, A. 1993. Towards an integrated logic of space, time, and motion. *IJCAI*.

Galton, A. 1997. Space, time, and movement. Stock, O., ed., *Spatial and temporal reasoning*. Kluwer.

Herskovits, A. 1986. *Language and Spatial Cognition: an Interdisci- plinary Study of the Prepositions in English*. Cambridge University Press.

Jackendoff, R. 1983. *Semantics and Cognition*. MIT.

Kipper, K., A. Korhonen, N. Ryant, and M. Palmer. 2006. Extensive classifications of english verbs. In *12th EURALEX*, Turin, Italy.

Mani, I., J. Hitzeman, and C. Clark. 2008. Annotating natural language geographic references. In *Workshop on Methodologies and Resources for Processing Spatial Language*. LREC'2008.

MITRE. 2007. Spatialml: Annotation scheme for marking spatial expressions in natural language. http://sourceforge.net/projects/spatialml/.

Muller, P.. 1998. A qualitative theory of motion based on spatio-temporal primitives. Cohn, A., L. Schubert, and S. Shapiro, eds., *KR'98: Principles of Knowledge Representation and Reasoning*, pages 131–141. Morgan Kaufmann.

Muller, P. 2002. Topological spatio-temporal reasoning and representation. *Computational Intelligence*, 18(3):420–450.

Pustejovsky, J. and M. Verhagen. 2007. Inferring spatio-temporal trajectories of entities from natural language documents. Tech. report, Brandeis U.

Pustejovsky, J., J. Castaño, R. Ingria, R. Saurí, R. Gaizauskas, A. Setzer, and G. Katz. 2003. Timeml: Robust specification of event and temporal expressions in text. In *IWCS-5, Fifth International Workshop on Computational Semantics*.

Pustejovsky, J. 1995. *The Generative Lexicon*. MIT.

Randell, D. A., Z. Cui, and A. G. Cohn. 1992. A spatial logic based on regions and connections. Kaufmann, M., ed., *3rd International Conference on Knowledge Representation and Reasoning*.

Vieu, L. 1991. *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.