[Review]

Interpreting Motion: Grounded Representations for Spatial Language


TORU SUZUKI
Yamagata University*

Keywords: computational linguistics, motion, qualitative reasoning, spatial prepositions, verb typology

1. Introduction

In everyday life, the concept of motion is ubiquitous, if not always obvious to our consciousness: when we experience or perceive our most common physical activities such as walking, running, dancing or swimming, motion comes at the center of our mental description of events such as these located in space and time. As a consequence, understanding how we conceptualize and verbalize motion becomes one of the most substantial issues for the study of human mind and communication that ranges across linguistics, formal semantics, computational linguistics, philosophy, cognitive science, geographical information systems and many other fields.

With unparalleled advances in modern technology, the design and implementation in particular of computer programs capable of automatically processing text corpora containing descriptions of the motion involved in human activities should have a significant practical consequence, allowing us to map from texts to data representations that can be of great value in our daily life.

Inderjeet Mani and James Pustejovsky, the two authors of the book under review, Interpreting Motion: Grounded Representations for Spatial Language, point out, however, that “while there has been a great deal of lin-

* I am grateful to Mark Irwin for his suggestions for stylistic improvements. Research for this review is supported in part by a Grant-in-Aid for Scientific Research (C) (No. 24520528).

© 2015 by the English Linguistic Society of Japan
linguistics research on the semantics of motion verbs as well as locative constructions, and considerable research on qualitative spatial reasoning, there has been little interdisciplinary effort on trying to connect these two fields in a systematic way” (pp. 5–6). Based on this perspective, the book aims at integrating the language of motion with ‘qualitative spatial reasoning,’ an area of Artificial Intelligence (AI) research on reasoning for problem solving. It is ‘qualitative’ rather than ‘quantitative’ because “[h]aving an artificial agent reason qualitatively allows for reasoning to be more efficient in some situations, since abstracting away from numerical details allows the agent to focus on more compact representations that isolate just the relevant information needed to solve a particular problem” (p. 7).

Their specific goals are two-fold: from a theoretical perspective, the book proposes a semantic theory of motion expressions in natural language that can be used for computation; from a pragmatic perspective, the authors try to present an integrated computational system that can automatically process natural language corpora containing descriptions of spatial, temporal, and motion-related information. In order to achieve these goals, the authors take an interdisciplinary approach combining the linguistic conceptualization of various motion expressions with the formal computational methods of qualitative reasoning, mapping one to the other in the context of natural language processing.

A caveat for the reader: although the book contains a lot of detailed discussion and illustration of the computational implementations of their theory, the present reviewer, being an ‘ordinary’ linguist of syntax and semantics, does not pretend to be able to follow every computational detail of the framework developed in the book. Therefore, the following review is largely focused on ‘linguistic’ areas of their framework.

2. Overview

The book consists of six chapters: odd-numbered chapters are mainly written by Mani, even-numbered chapters by Pustejovsky, with the final chapter (Chapter 6) being a contribution from both authors.

In Chapter 1 the book sets forth the basic requirements for the interdisciplinary approach to be adopted in the book: “the semantic representations need to be expressive enough for natural languages, but also must be amenable to inference methods that can be used in practical systems” (p. 4). Their essential desiderata are stated more specifically as follows:
1. The semantic representations need to be expressive enough for natural languages, but also must be amenable to inference methods that can be used in practical systems.
2. The semantic theory must be denotational/truth-conditional.
3. The semantic analysis must be compositional.
4. The representations have to support qualitative reasoning.
5. The systems must be accurate and efficient enough to support practical applications. (pp. 10–11)

In light of their own desiderata, the authors provide a critical assessment of previous and current linguistic approaches to the semantics of spatial prepositions and motion verbs, including procedural semantics (Miller and Johnson-Laird (1976)), cognitive linguistic approaches (Herskovits (1986), Jackendoff (1983, 1990), Langacker (1987)) and vector representations (Zwarts (2003), Zwarts and Winter (2000)), FrameNet (Baker et al. (2003)), among others. On the semantics of spatial prepositions, the authors summarize their review as “[w]hile compositional treatments of prepositional meaning have flourished, the question of what underlying spatial primitives to rely on has not thus far been tied to those available in qualitative reasoning systems” (p. 18); as for the semantics of motion verbs, they virtually reject most of the previous approaches (apart from Talmy (1985, 2000)), mainly for their lack of explicitness in formalization and amenability to spatial reasoning, although they do cite some theoretical insights they are willing to adopt. For example, they point out the shortcoming that the distinction between the motion verbs enter and arrive is not adequately explicated by the above theories.

The introductory part of this book also underlines the question of compositionality, namely how the meanings of various constituents in an expression are specified and composed without undue proliferation of lexical senses. In order to achieve the fine-grained semantic classification of spatial meaning and the compositionality requirement, the Generative Lexicon (GL: Pustejovsky (1995)) is incorporated in their framework. An illustrative example is given with the expression “the coffee in the cup,” where the meaning of “in” can be specified without proliferating senses. The preposition “in” has an underspecified notion of containment (inside a container), the noun “cup” is a drinking instrument (an open container), and “coffee” has a noun sense of liquid material. To compose these elements to yield the meaning of “coffee in the cup,” the liquid has to be contained in the container such that its convex hull (the boundary formed by the minimal convex set containing a region as a set of points) is required to be inside
the container. This is “achieved within a compositional semantics using GL (based on notions of coercion and co-composition), via an axiom of world knowledge” (p. 26).

Spatial meaning is also derived from elements other than spatial prepositions and motion verbs. In order to deal with the fact that “many objects have an ‘intrinsic’ front, bottom, etc., that is highly culture- and use-dependent” (p. 68), for example, the Qualia structure (‘telic roles’) in GL is exploited to represent the intrinsic frame of reference (orientations) of entities, in particular, artifacts such as books, TVs, planes, etc. The canonical orientation of such entities largely depends on functional criteria, on how the object is typically used or presented. In GL, the ‘telic roles’ indicate how artifacts are used, and as part of that use, their inherent orientations can be enumerated as components of the lexical entry. In their treatment of spatial meaning, therefore, the Qualia structure constitutes an indispensable part of adequately representing semantic contributions from nominal expressions with respect to orientations based on the figure/ground distinction.

In Chapter 2, the authors discuss a theoretical framework in which motion events are characterized in natural language. In order to compose denotational event semantics together with spatial expressions, they assume two types of spatial meaning parameter: (i) topological relations between two objects defined by the relation of connectedness enriched with notions of orientation and distance, and (ii) semantic parameters for description of motion events introduced by the lexical argument structure of verbs and other syntactic categories involving motion.

The topological relations between two objects A and B defined by connectedness (the Region Connection Calculus 8 (RCC8) adopted from Randell et al. (1992)) consist of eight jointly exhaustive and pairwise disjoint relations, as shown in (1).

(1) a. Disconnected (DC): A and B do not touch each other.
b. Externally Connected (EC): A and B touch each other at their boundaries.
c. Partial Overlap (PO): A and B overlap each other in Euclidean space.
d. Equal (EQ): A and B occupy the exact same Euclidean space.
e. Tangential Proper Part (TPP): A is inside B and touches the boundary of B.
f. Non-tangential Proper Part (NTPP): A is inside B and does not touch the boundary of B.
g. Tangential Proper Part Inverse (TPPi): B is inside A and touches the boundary of A.

h. Non-tangential Proper Part Inverse (NTPPi): B is inside A and does not touch the boundary of A.

While it is shown that systems of qualitative spatial reasoning equipped with RCC-8 can adequately handle static information about space, a further set of parameters is called for to properly represent the dynamic aspects of motion events.

The semantic parameters of motion (basic motion frame) have essentially been adopted from the seminal work on cognitive semantics of motion by Talmy (1985, 2000) and are:

(2) a. The event involved in the change of location.
   b. The figure undergoing movement.
   c. The path traversed through the motion.
   d. A distinguished region of the path, the ground.
   e. The manner in which the change of location is carried out.
   f. The medium through which the motion takes place.

Here, departing from the original characterization by Talmy, the authors suggest manner (in (2e)) be treated not as a primitive characteristic of motion, but instead as an elaboration of any of the remaining five parameters. “The characterization of manner is merely the modification of other aspects of the motion frame” (p. 49). Hence, manner properties may be distributed and heterogeneous in nature. In this view, the modification of any of the other elements in the motion frame will be considered a manner specialization.

In addition to the basic motion frame, a well-known distinction of path predicates (movement with distinguished location) and manner-of-motion predicates (movement without distinguished location) is introduced: path verbs (‘path presupposing predicates’) are those verbs that presuppose a specific path for the moving object (the figure), along with a possible distinguished point or region on this path (the ground) which the figure is moving toward or away from (e.g. arrive, leave, ascend, descend, approach); manner of motion verbs (‘path creating predicates’) are considered subtypes of an atomic predicate for motion move, which takes as its arguments the figure and the tracing of the movement by the object in motion, which is a path (e.g. bike, drive, walk, fly).

Chapter 3 is an examination of the mapping from spatial and temporal expressions to computational formalism for qualitative reasoning. In the main, the authors discuss the two geometric notions of space, namely topology and orientation, while also pointing out that non-geometric aspects of
meaning, such as function, do play a role in interpretation. To obtain the necessary information for representing entities (in particular, artifacts) with inherent orientations and/or functions, the reader is again invited to tap into the theory of GL.

In Chapter 4, the authors demonstrate, as an application of the methods discussed so far, how the two linguistic strategies for encoding motion ('path constructions' using a path verb (*enter*, *arrive*) with a manner adjunct and 'manner-of-motion constructions' using a manner-of-motion verb (*run*, *bike*) with a path adjunct) are modeled as well as how prepositional, noun, and verb meanings are integrated compositionally within a new first-order Dynamic Interval Temporal Logic (DITL, Pustejovsky and Moszkowics (2011)). Two different types of compositional constructions with a manner verb (e.g. *John biked in the morning to Agua Azul*) and a path verb (e.g. *John left Ocosingo this afternoon by foot*) are schematically illustrated in Figures 1 and 2 respectively:

**Figure 1** Manner verb + Path PP: *bike to Agua Azul* (p. 102)
Chapter 5 discusses the methodology of linguistic annotation, examining specific algorithms for linguistic analyses of motion that derive information from text corpora and eventually allow for automatic text-to-sketch mapping. In particular, ISO-Space (Pustejovsky et al. (2011)) is highly evaluated because of its rich representation of paths and well-defined distinction between manner and path verbs as well as various subclasses of motion events. All in all, the authors conclude that “[t]he capabilities for assembling automatic motion tracking from natural language narratives are well underway” (p. 127).

In the final chapter (Chapter 6), some of the practical applications of the theory of representing the meanings of motion verbs and spatial expressions in terms of qualitative reasoning are demonstrated. These are done so in terms of route navigation, mapping of travel narratives, multimedia tagging (of static images, audio, and video), question-answering, and so on. The authors conclude the book with some brief remarks on future challenges, including issues of metaphorical motion (including ‘fictive motion’), cross-linguistic variation, 3D spatial representations, the tractability problem (efficiency), integration of reasoning and extraction, and trainability in data preparation.

3. Assessment

Some potentially controversial points can be found in the authors’ self-
imposed desiderata. As mentioned earlier, the authors criticize most of the previous cognitive semantic approaches to the language of space and motion for their lack of compositionality and denotationality for further computation. Although methodologically comprehensible in terms of its specific computational purpose, the utility of formalization in linguistics itself is a rather debatable issue: how far and adequately formalization can capture our linguistic understanding of motion, we still do not know. In fact, in other computational linguistic approaches that seem to share most of the issues under discussion, a more liberal stance is taken in which a compositional semantic level is more flexibly and efficiently separated from a non-compositional ‘conceptual’ level (Bateman et al. (2010)). In a similar vein, with emphasis on their formal semantic approach, the authors explicitly omit pragmatic aspects of our understanding of motion and space: “Questions of pragmatics, which of course are key to the understanding of language in context, are not addressed” (p. 27). In this regard, the reader is encouraged to put the book in a comparative perspective along other more eclectic approaches across the relevant fields.

In describing motion events in linguistic terms, the authors draw particularly upon some of the key insights of Talmy (1985, 2000), including notions such as figure, ground, path and the typological distinction in motion expressions between ‘satellite-framing’ and ‘verb-framing.’ Although each of the former primitive notions is considered to constitute a semantic component in their framework, it should be noted that originally Talmy (1985, 2000) himself intends those notions to be ‘conceptual’ rather than ‘semantic.’

Talmy’s dichotomy of ‘satellite-framing’ and ‘verb-framing’ is extended to two different classes of logical predicates, ‘action-based predicates’ and ‘location-based predicates,’ differentiating “movement without distinguished location” (manner of motion verbs) and “movement with distinguished location” (path verbs). Notably, however, the dichotomy itself is currently under extensive reexamination (Slobin (2004, 2006), Talmy (2009) among others), and although the authors provisionally acknowledge the recent shift in typology, including a third category ‘equipollent-framing’ (p. 9), future research framed in the dual system (manner of motion verbs vs. path verbs) might be affected in some way.

Lastly, one occasionally feels somewhat confused to find the terms ‘motion’ and ‘movement’ being sometimes used almost interchangeably, although, in common linguistic terminology, the former refers to an event of ‘displacement (translational motion)’ and/or ‘movement,’ while the latter is
normally used to refer to the manner of ‘displacement.’ A clearer distinction in motion terminology might have been helpful in this book.

4. Conclusion

The book under review is highly informative as an extensive overview of the issues concerning linguistic studies of space and motion. It puts special emphasis on the computational methodology of the automated mapping of texts to data representation in qualitative reasoning. On the other hand, it may not be so easily readable for general linguists, undertrained or untrained in computational linguistics, or for that matter, AI theories (including the reviewer). Still, it allows us to have a firsthand glimpse of the state-of-the-art advances in the computational linguistics of motion. Here, various linguistic analyses of motion and spatial events can be tested empirically, both in terms of their applicability in the face of real-life language practice, and in terms of their explicitness and relevance for representing the information required for further computation and reasoning.

REFERENCES

Pustejovsky, James, Jessica Moszkowicz and Marc Verhagen (2011) “ISO-Space: The Annotation of Spatial Information in Language,” *Proceedings of ISA-6: ACL-


[received December 20, 2014, revised and accepted May 9, 2015]

Faculty of Literature and Social Sciences
Yamagata University
1–4–12 Kojirakawa, Yamagata, Yamagata 990–8560
e-mail: torus@human.kj.yamagata-u.ac.jp