

A COMPUTATIONAL VIEW OF THE COGNITIVE SEMANTICS OF SPATIAL PREPOSITIONS*

Patrick Olivier

*Centre for Intelligent Systems
University of Wales*

Aberystwyth

Dyfed, SY23 3DB, UK

Internet: plo@aber.ac.uk

Jun-ichi Tsujii

*Centre for Computational Linguistics
University of Manchester*

Institute of Science and Technology

Manchester, M60 1QD, UK

Internet: tsujii@ccl.umist.ac.uk

Abstract

This paper outlines the linguistic semantic commitments underlying an application which automatically constructs depictions of verbal spatial descriptions. Our approach draws on the ideational view of linguistic semantics developed by Ronald Langacker in his theory of Cognitive Grammar, and the conceptual representation of physical objects from the two-level semantics of Bierwisch and Lang. In particular the dimensions of the process of *conventional imagery* are used as a metric for the design of our own conceptual representation.

INTRODUCTION

An increased interest in the semantics of spatial language has accompanied the recent rise in popularity of cognitive linguistics (see [Rudzka-Ostyn1988]), yet computational approaches are thin on the ground. This can in part be accounted for by the rather descriptive and unformalized nature of the theories developed, but is more likely due to the adoption of an ideational view of linguistic meaning which, it seems, is an anathema to computational linguists. In this paper we take a serious, if informal, look at Ronald Langacker's theory of Cognitive Grammar [Langacker1987], [Langacker1988a], [Langacker1988b], more specifically its commitment to conceptualization and the use of conventional imagery.

The first section of this paper introduces the semantics of projective prepositions (eg. "in front of", "behind", "left of", "right of"), illustrating that these seemingly simple predicates are surprisingly complex and ambiguous. In the light of this discovery the following sections consider Langacker's view of linguistic meaning, and the design of a conceptual representation for spatial prepositions motivated by the consideration of the various

*This research was kindly funded by the Matsushita Electric Industrial Company Limited.

dimensions of conventional imagery. The representation has been implemented for English spatial descriptions and after demonstrating its utility for the automatic depiction of verbal descriptions, we finally contrast our approach against previous attempts.

THE SEMANTICS OF PROJECTIVE PREPOSITIONS

In this section we characterize the components of the spatial meaning of projective prepositions that have motivated our interest in cognitive linguistic approaches. Throughout, the decoding problem, that is, generating adequate meanings for a locative expression in a particular situation, is our benchmark for representational adequacy.

The spatial meaning of a projective prepositional predication (eg. "the chair is in front of the desk") can include: a constraint on the proximity of the located (LO) (eg. "the chair") and reference objects (RO) (eg. "the desk"); a directional constraint on the LO relative to the RO; and a relative orientation between the speaker, LO and RO. Constraints are of an intrinsically fuzzy nature such that different relative positions and orientations of the speaker, RO and LO satisfy the predication to different degrees, and combinations of constraints on the RO and LO originating from different predications must be readily accommodated.

PROXIMITY CONSTRAINTS

Projective prepositions necessarily place a constraint on the proximity of the located object and the reference object. Predications such as "the chair is in front of the desk" constrain the "desk" and "chair", to some degree, to be proximal to each other. Conversely projective prepositions such as "away from" predicate a distal relationship between the located and reference object. The degree of the proximity expressed in any projective prepositional predication varies accord-

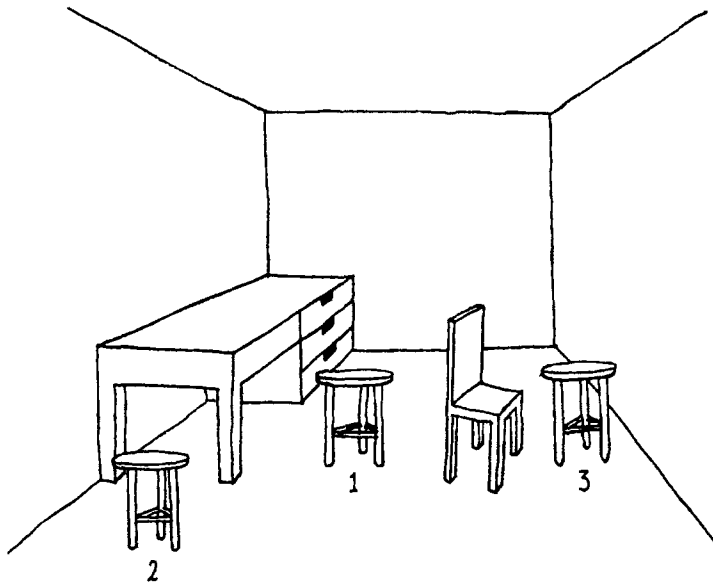


Figure 1: Intrinsic, deictic and extrinsic uses of "in front of"

ing to a number of considerations including: the spatial context (the spatial extent and content of the scene described); and the absolute and relative sizes of the LO and RO (eg. a car that is "left of" a lorry is typically less proximal than an apple and orange similarly described).

DIRECTIONAL CONSTRAINTS

In addition to the constraint on the proximity of the LO and RO, projective prepositions place a constraint on the position of the LO relative to a particular side of the RO. In the case of the intrinsic interpretation (see section) of a predication such as "the stool is in front of the desk", the "stool" is located in some region of the space defined by the half-plane that is the intrinsic front of the "desk". Intuitively, the closer the "stool" is to the region of space defined by the projection of the desk's dimensions into this space, the more the spatial arrangement conforms to the prototypical interpretation of the predication.

REFERENCE FRAMES

Intrinsic, deictic and extrinsic interpretations of projective prepositions differ according to the reference frame with respect to which the directional constraint is characterized [Retz-Schmidt1988]. Figure 1 is an example of a scene that might give rise to predications which invoke each of these reference frames.

INTRINSIC In the intrinsic case the reference frame is centered at the RO and adopts the intrinsic orientations of the RO. Thus a LO is deemed to be "in front of" the RO under an intrinsic reading if it is located in the direction defined by the vector that is the half-plane of the *front* of the RO. In figure 1 stool number 1 is intrinsically "in front of the desk".

DEICTIC The reference frame for a deictic interpretation is centered at the speaker and adopts the speaker's orientation; deictic readings can be invoked explicitly with qualifications such as "from where we are standing"; when the RO has no intrinsic or extrinsic sidedness relating to the preposition used; or when intrinsic or extrinsic interpretations are ruled out on other grounds (eg. the impossibility of spatially arranging the objects as required by the interpretation). In figure 1 stool number 2 is deictically "in front of the desk".

EXTRINSIC Extrinsic readings can occur when the RO has no intrinsic sides relating to the locative preposition (eg. for objects such as trees) but is in close proximity to another object that is strongly sided (eg. such as a house); in which case the reference frame capturing the intrinsic orientations of the stronger sided object can be adopted by the RO. Referring to figure 1 the chair is extrinsically "in front of stool number 3"; here the stool has inherited an extrinsic front from the right wall.

INTERACTING CONSTRAINTS

Typically an object is located with respect to more than one RO by the means of multiple spatial predications. This places a requirement of on the meaning representation of spatial predications that they must be capable of being easily combined, to give rise to a cumulative meaning.

COGNITIVE GRAMMAR AND LINGUISTIC MEANING

Cognitive grammar is comprised of five basic claims as to the composition of linguistic meaning, following [Langacker1988b] these are:

1. Meaning reduces to conceptualization.
2. Polysemy is the norm and can be adequately accommodated by representing the meaning a lexical item as a network of senses related by categorizing relationships of schematicity or extension.
3. Semantic structures are characterized relative to cognitive domains. Domains are hierarchically

organized in terms of conceptual complexity, where the characterization of a concept at one level can draw on lower level concepts. While there need not necessarily be any conceptual primitives, the lowest level domains are termed basic domains and include our experience of time, space, color etc.

4. A semantic structure derives its value through the imposition of a "profile" upon a "base".
5. Semantic structures incorporate conventional "imagery", our ability to construe the same informational content in different ways.

That meaning reduces to conceptualization (thesis 1), is characterized relative to cognitive domains (thesis 3), and incorporates conventional imagery (thesis 5) runs in stark contrast to the heavy emphasis placed on truth conditions and formalization by current computational linguistic approaches. We have attempted to tackle the informality of this ideational view of meaning, by addressing one particular basic cognitive domain, that of oriented three-dimensional space, and implement a restricted version of Langacker's process of conceptualization by means of conventional imagery. To verify the utility of the resulting conceptualization, we use the interpretations of spatial expressions so generated (the resulting images), to automatically construct a depictions of the scene.

Theses 2, that prototypes should replace traditional objective categories, lies at the very heart of cognitive semantics [Taylor1989], and though it is widely accepted as true for semantic and most other linguistic categories, prototype theory is not conducive to rigorous formalization and has consequently been ignored by mainstream computational linguistics. Likewise our concern is with meaning variations that originate from different construals of the same information in the process of conventional imagery (thesis 5).

IMAGERY AND ITS IMPLEMENTATION

This special technical use of *imagery* (not to be confused with the psychological term meaning the formation and manipulation mental images) refers to "our amazing mental ability to "structure" or "construe" a conceived situation in many alternate ways" [Langacker1988b], as opposed to traditional semantic approaches whose concern is with informational content alone. Thus "every conception reflects some particular construal of its content". Langacker identifies six important dimensions of imagery; in our semantic analysis of spatial expressions we are interested in just three of these:

1. level of specificity

2. scale and scope of predication
3. perspective

The remainder of this section is a characterization of each of these dimensions and the consequences that their consideration has with respect to the design of a conceptual representation for spatial expressions.

REPRESENTING 3-D SPACE

The basic cognitive domain relative to which the spatial meaning of projective prepositions is characterized, is structured three-dimensional space. In our system space is represented using an orthogonal axis system we refer to as the DCS (Domain Coordinate System). In the process of image construction conceptual objects will be constrained to locations described relative to the DCS. The DCS mirrors the speaker's perceptual assignment of axes to a scene, the *x*-axis extends from deictic left to deictic right, the *y*-axis from deictic front to deictic back, and the *z*-axis extends vertically.

LEVEL OF SPECIFICITY

The level of specificity of conventional imagery addresses the issue of the degree of detail with which an entity is characterized. Specificity has already been mentioned in connection with the construction of the network of polysemous senses of a lexical item; on the other hand, concerning different lexical items, we can readily identify different spatial predications that are schematic with respect to each other. Consider the sentences below.

- (a) The chair is near the desk.
- (b) The chair is in front of the desk.
- (c) The chair is facing the desk.

Sentence (a) simply predicates proximity; (b) predicates both proximity and a positioning of the LO relative to a particular side of the RO¹; lastly (c) predicates proximity and a relative positioning of the LO with respect to the RO, with the additional anti-alignment of the *front* face normals of the two objects.

Schematic contrast dictates the minimum degree of detail we must maintain in our computational representation of the conceptual reference and located objects. In sentences (a) the objects can be thought of as structureless points; in (b) the representation of the RO must incorporate the notion of sidedness; and in (c) both the RO and LO are sided. We borrow Lang's conceptual representation of objects

¹The issue of which side of the reference object the located object is positioned with respect to is addressed as a consequence of the perspective dimension of conventional imagery

termed object schemata [Lang1993], constructed within Bierwisch's and Lang's the two-level semantics [Bierwisch and Lang1989]. The object schema for a desk is:

a	max	b	vert	c	across
a1	i-left	b1	i-bottom	c1	i-front
a2	i-right	b2	i-top	c2	i-back

In this first schema **a**, **b** and **c** label three orthogonal axes centered at the object, each of which can be instantiated by one or more dimensional assignment parameters (DAPs)²; **a1-a2**, **b1-b2** and **c1-c2** are corresponding half-axes. Each half axis is labelled either **nil** or with an intrinsic side (eg. **i-front**). This representation is augmented with both a three-dimensional Cartesian coordinate which when assigned locates the conceptual schema relative to the DCS; and the values of the default extents for the object type along the axes **a**, **b** and **c**.

Imagery implies an *imager*, that is, the image exists in and with respect cognitive world of the speaker (by default) and this necessarily has important consequences. With respect to spatial language, issues pertaining to perspective, that is taking account of the imager, include the speaker's vantage point and orientation.

ORIENTATION

The interpretation of some spatial expressions is dependent on assumptions as to the speaker's orientation with respect to the objects in the scene (eg. whether A is "to the left of" B in a scene, is dependent on the orientation of the speaker/viewer); other expressions are orientation independent such as "above" and "below" which implicitly refer to the downward pull of gravity (although in space verticality is speaker dependent).

When an object schemata is characterized relative to the DCS it is both assigned a Cartesian position (as we show later), and its half-axes are assigned deictic sides according to their relative orientation with the observer. For example if a desk is positioned "against the left wall" as in figure 1 this would result an instantiated conceptual schema for the "desk" of:

a	max	b	vert	c	across
a1	i-left	b1	i-bottom	c1	i-front
	d-front		d-bottom		d-right
a2	i-right	b2	i-top	c2	i-back
	d-back		d-top		d-left

²DAPs are not of direct interest here although they are fundamental to the process of dimensional designation and an important where dimensional assignment might result in a reorientation of the conceptual object (eg. "the pole is high").

Here **a1** is the intrinsic left side but the deictic front of the desk.

VANTAGE POINT

The speaker's vantage point is another factor that determines the interpretation of spatial expressions in a scene. The notions of deictic and intrinsic interpretations of projective prepositions can be accounted for purely by recognizing that in each the speaker adopts a different vantage point. For deictic interpretations the vantage point is the speaker's actual position. The vantage point for intrinsic interpretations is the functionally relevant position with respect to a reference object, for example, "left of the desk" under the intrinsic interpretation uses a vantage point that is directly in front of the desk (the typical configuration when a human uses a desk).

The meaning of a projective preposition is conceptually represented as a spatial constraint on the conceptual schema of the located object which extends out from a particular side of a reference object, the precise nature of which we describe in the next subsection. In our system the lexicalized constraint is of the form of a two place predicate: $\langle \text{zoneprox } X:\text{side } Y \rangle$ Where **X** is the reference object and **Y** the located object. The parameter **side** depends on the preposition. Thus the schematicity we observed in section is explicitly represented:

- (a) Y is near X.
 $\langle \text{zoneprox } X \ Y \rangle$
Proximity constraint between X and Y.
- (b) Y is in front of X.
 $\langle \text{zoneprox } X:\text{front } Y \rangle$
Proximity and alignment of Y with front of X
- (c) Y is facing X.
 $\langle \text{zoneprox } X:\text{front } Y:\text{back} \rangle$
Proximity, alignment and specific "facing" orientation

SCOPE OF PREDICATION

Scope refers to exactly how much of a cognitive domain is included in the characterization. Minimally, the scope of an image for "next to" must encompass at least the reference and subject objects and some region of space separating them. We implement the spirit of this concept by realising the lexicalized constraint for a projective preposition as a potential field fixed at the reference object's position in the DCS³. The proximity and directional nature of the constraint $\langle \text{zoneprox } .. \rangle$ is captured using a potential field P_x , where:

$$d_x = (x - x_0) \quad (1)$$

³This technique is borrowed from robot manipulator path-planning [Khatib1986]

$$d_y = (y - y_0) \quad (2)$$

$$P_{\mathcal{X}} = P_{prox, \mathcal{X}} + P_{dir, \mathcal{X}} \quad (3)$$

$$P_{prox, \mathcal{X}} = \frac{K_{prox, \mathcal{X}}}{2} (\sqrt{d_x^2 + d_y^2} - L_{prox, \mathcal{X}})^2 \quad (4)$$

$$P_{dir, \mathcal{X}} = \frac{K_{dir, \mathcal{X}}}{2} d_x^2 \quad (5)$$

Here the x -axis points direction of the half-axis of the particular side of the reference axis in the DCS; and in the case of "in front of" y is the perpendicular direction in the horizontal plane; (x_0, y_0) is the Cartesian coordinate of the reference object in the DCS, and lower the value of $P_{\mathcal{X}}$ for a location (x, y) for the located object the *better the spatial constraint is satisfied*. The minimum for the field can be quickly computed using gradual approximation [Yamada et al.1988]. The values of $K_{prox, \mathcal{X}}$, $L_{prox, \mathcal{X}}$, and $K_{dir, \mathcal{X}}$ are dependent on the located and reference objects and are set on the basis of *scale* considerations (see). Multiple spatial predications over an object is simply accommodated within the potential field model by linear addition of component fields.

SCALE OF PREDICATION

The concept of the scale relates to the object dependency of the degree of proximity and directional constraint afforded by a preposition: where "X is left of Y", and X and Y are houses, then the meaning of this predication would contrast with its meaning if X and Y were pieces of fruit. The concept of proximity and directional constraint predicated by "left of" is apparent in both cases, what differs is the scale relative to which it is characterized.

Scale effects are realised in the mechanism by which the constants of the potential field are set. For the potential field $P_{\mathcal{X}}$, the effect of the constants on the nature of the constraint are:

1. $K_{prox, \mathcal{X}}$
Proportional to range of the possible separations of X and Y that would still satisfy the predication.
2. $L_{prox, \mathcal{X}}$
The default separation of X and Y.
3. $K_{dir, \mathcal{X}}$
Proportional to the range of directions that would still satisfy the predication.

Thus for a reference object that is a house $K_{prox, \mathcal{X}}$, $L_{prox, \mathcal{X}}$, $K_{dir, \mathcal{X}}$ must all be considerably greater than for a piece of fruit. The precise values can only reasonably set as a result of some experimental investigation, currently $K_{prox, \mathcal{X}}$ and

$L_{prox, \mathcal{X}}$ are linearly dependent on the sum of the extents of the reference and subject objects in the direction of spatial alignment; and $K_{dir, \mathcal{X}}$ on the perpendicular extent of the reference object in the plane of the constraint.

GENERATING DEPICTIONS

After using gradual approximation to find the position of the minimum in the potential fields representing the spatial predications over a particular object, this point can be regarded as a probable interpretation. By tying each conceptual object to a graphical model, and interpreting the DCS as the viewer's perceptual axis system, conceptual interpretations can be rendered as scene depictions. Figure 2 illustrates one depiction of the cumulative interpretation of the following verbal description, in which all projective prepositions are viewed intrinsically⁴.

"I am in a room. Against the left wall is a long desk. Against the back wall is a short desk. In front of the long desk is a chair. Another chair is to the left of the long desk. The chair in front of the desk is near the short desk."

OTHER APPROACHES AND CLOSING REMARKS

Nearly all the work in recent years on computing the meanings of spatial prepositions stem from the prototype semantics of either Herskovits [Herskovits1985], [Herskovits1986] or Talmy [Talmy1983]. Schirra [Schirra and Stopp1993] adopts Herskovits' notion of a core meaning, and implements this as a typicality field. The ability to sum fields of different predications satisfies the compositionality requirement. Yet representational poverty exists with respect to the spatial and perceptual characteristics of the objects, as while directionality and proximity constraints are adequately captured for the intrinsic reference frame and set of objects, variation in the degree of constraint (for example, depending on the size of the reference object) and the potential for ambiguity arising from interpretations with respect to different reference frames are not accounted for.

Underlying Kalita's work [Kalita and Badler1991] is a conceptualization of the space around a reference object as six

⁴Natural language sentences are parsed to three branch quantifiers using a prolog DCG grammar, the logical predicates are the input to the cognitive semantic processor, the resulting conceptual representations are converted to depictions in by the depiction module . The cognitive semantic processor and the depiction module are implemented in Smalltalk/Objectworks

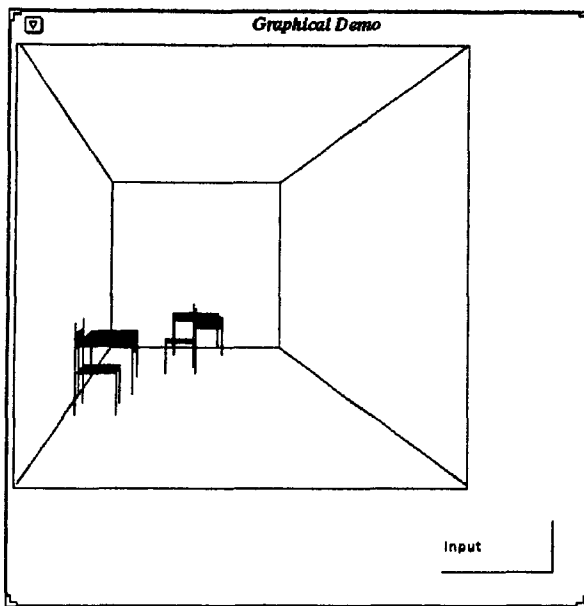


Figure 2: Computer generated depiction of a verbal description

orthogonal rectangular projected regions (based upon an enclosing cuboid idealization of the object) due to Douglas [Douglas and Novick1987]. Using this model and following Talmy's work, the semantics of projective prepositions are lexicalized as geometric-relation schemas. Reference frame ambiguity is not addressed; directionality is too tightly restricted to one of the six rectangular regions, and proximity constraint is left to the "underlying constraint satisfaction techniques and the use of a weight slot in the template for constraint representation".

Within the framework of the LILOG project [Maienborn1991] Ewald Lang implemented the two-level approach to the semantics of dimensional adjectives in which the perceptual and dimensional properties of objects are conceptually represented as object schemata [Bierwisch and Lang1989]. Further developed for projective spatial predications, Lang's object schemata are capable of distinguishing deictic and intrinsic readings, though without explicit reference to a quantitative space (ie. actual scenes and observers) as in the case of Schirra and Kalita.

Our system represents a first attempt, and very highly specialized implementation, of the conventional imagery process that is a component of the cognitive grammarian's view of linguistic semantics. Its performance, in terms of generating all possible interpretations, and the quality of the interpretations constitutes a significant advance

on previous approaches.

References

- [Bierwisch and Lang1989] M Bierwisch and E Lang. 1989. *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*. Springer-Verlag, Berlin Heidelberg New York.
- [Douglas and Novick1987] S Douglas and D Novick. 1987. Consistency and variance in spatial reference. In *Proceedings of the Ninth Annual Cognitive Science Society Meeting*, pages 417-426.
- [Herskovits1985] A Herskovits. 1985. Semantics and pragmatics of locative expressions. *Cognitive Science*, 9:341-378.
- [Herskovits1986] A Herskovits. 1986. *Language and spatial cognition — an interdisciplinary study of the prepositions in English*. Cambridge University Press, Cambridge (UK).
- [Kalita and Badler1991] J Kalita and B Badler. 1991. Interpreting prepositions physically. In *Proceedings AAAI-91*, pages 105-110.
- [Khatib1986] O Khatib. 1986. Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research*, 5(1):90-98.
- [Lang1993] E Lang. 1993. A two-level approach to projective prepositions. In C Zelinsky-Wibbelt, editor, *The semantics of prepositions: from mental processing to Natural Language processing*. Mouton de Gruyter, Berlin.
- [Langacker1987] R W Langacker. 1987. *Foundations of Cognitive Grammar, Volume I, Theoretical Prerequisites*. Stanford University Press, Stanford.
- [Langacker1988a] R W Langacker. 1988a. An overview of cognitive grammar. In B Rudzka-Ostyn, editor, *Topics in Cognitive Linguistics*, pages 3-48. Benjamins, Amsterdam-Philadelphia.
- [Langacker1988b] R W Langacker. 1988b. A view of linguistic semantics. In B Rudzka-Ostyn, editor, *Topics in Cognitive Linguistics*, pages 49-90. Benjamins, Amsterdam-Philadelphia.
- [Maienborn1991] J R Maienborn. 1991. Processing spatial knowledge in lilog. IWBS Report 157, IBM Germany.
- [Retz-Schmidt1988] G Retz-Schmidt. 1988. Various views on spatial prepositions. *AI Magazine*, 9(2):95-105.
- [Rudzka-Ostyn1988] B Rudzka-Ostyn, editor. 1988. *Topics in Cognitive Linguistics*. Benjamins, Amsterdam-Philadelphia.

- [Schirra and Stopp1993] J R J Schirra and E Stopp. 1993. Antlima — a listener model with mental images. In *Proceedings of IJCAI*, pages 175–180.
- [Talmy1983] L Talmy. 1983. How language structures space. In H Pick and L Acredolo, editors, *Spatial Orientation: Theory, Research, and Application*, pages 225–282. Plenum Press, New York.
- [Taylor1989] J R Taylor. 1989. *Linguistic categorization: prototypes in linguistic theory*. Oxford University Press, Oxford.
- [Yamada et al.1988] A Yamada, T Nishida, and S Doshita. 1988. Figuring out most plausible interpretation from spatial descriptions. In *Proceedings of the 12th International Conference on Computational Linguistics*, pages 764–769.