Proceedings of the Annual Meeting of the Cognitive Science Society

Consistency and Variation in Spatial Reference

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 9(0)

Authors

Douglas, Sarah A.
Novick, David G.
Tomlin, Russell S.

Publication Date

1987

Consistency and Variation in Spatial Reference [1]

Sarah A. Douglas
David G. Novick
Computer and Information Science Department
University of Oregon

Russell S. Tomlin
Linguistics Department
University of Oregon

Abstract

Modeling the meaning and use of linguistic expressions describing spatial relationships holding between a target object and a landmark object requires an understanding of both the consistency and variation in human performance in this area. Previous research [Herskovits 1985] attempts to account for some of this variation in terms of the angular deviation holding among objects in the visual display. This approach is shown to fail to account for the full range of human variation in performance, and a specific alternative algorithm is offered which is grounded in task variability and the notions of corridor and centroid. The significance to this algorithm of task variation, of the separation of semantic from pragmatic issues, and of the role of function and structure is discussed.

Keywords: spatial relations, natural language, reference.

1 Introduction

There is a growing body of literature in cognitive science which deals with the cognition of spatial relations. The aim of this research is to discover the principles which underlie the appropriate use and comprehension of expressions concerning spatial relationships among objects. However, current approaches to the problem of the language of spatial relations typically face difficulties in explaining why there seem to be so many different spatial relations with similar descriptions. Why do minor variations in physical relationships seem to give rise to major differences in language used to describe them, when in other cases major physical variations result in consistent linguistic descriptions? The answer to this question comes from the recognition that current theories of spatial relations tend to oversimplify the range of human linguistic performance that must be accounted for, and tend to conflate semantic concerns with pragmatic ones. In this paper, we show that accounting for these factors can explain one of the problems explicitly recognized as unaccountable in earlier work. We demonstrate that a more complex theoretical treatment of the grammar of spatial relations permits a good account of cases problematic for Herskovits [1985], and provides a practical algorithm for use in computational systems.

Herskovits [1985] looked at spatial relations in terms of "ideal relations"–what she might have called prototypes, except for the baggage of controversy over meaning which follows prototypes. Herskovits built on the work of, among others, Talmy [1983], who observed that in the enormous set of spatial relations among objects, only a relatively small number were lexicalized. Herskovits sought to explain lexical choice in locative descriptions. This paper expands on one aspect of her work–the ability of speakers and hearers to produce and accept apt locative descriptions where the relation between the referents deviates from the "prototype." Herskovits described this variation in angular terms:

---

[1] This research is supported by FIPSE grant #84.116C.

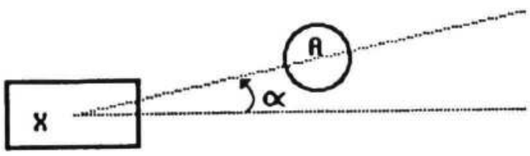


Figure 1: Is A ***directly to the right*** **of X?**

> There is a certain tolerance for deviation from truth of the ideal meaning, (or from the truth of the transformed ideal meaning when such a transformation has taken place.). I am concerned here with gradual deviations measurable in terms of an angle at a distance. For example, I want to ask: how far apart can two objects be so that one can say one is *at* the other? . . . Or consider [Figure 1]. How close to the right axis must an object A be, for:
>
> *A is directly to the right of X.*
>
> to be true? [Herskovits, 1985: 366]

Herskovits identifies indeterminacy arising from the nature of the objects and accuracy of perception, but argues that allowable variation (which she calls "tolerance"), depends chiefly on relevance of the relation to the discourse. In assessing the potential effects of these factors, she relies on angular deviation as the measure of variation. She then adds

> But often, the tolerance reflects an accumulation of practices, of interactions with the objects, making prediction impossible. In our constant intercourse with the objects in our world, we have integrated into our knowledge strategies that allow us to count or discount some fact according to context; what those strategies are is still very much a mystery. Yet, tolerance is one direction in which the search for systematicity could proceed. [Herskovits, 1985: 367]

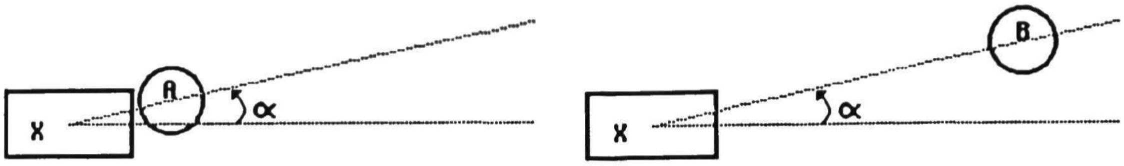
We believe that the mystery of acceptance of deviation from prototypical relations can be explained. The solution is in two parts. First, we show that angular differences are not the most useful measure of deviation. Second, we develop a plausible algorithm for description of spatial relations which inherently accounts for consistency and variation.

2 Acceptable Sources of Variation

Given the need to establish the basic semantics of spatial relations for *right of, left of, above, below*, and *between*, for the purpose of building an ICAI tutor for beginning second language instruction, we looked originally at the linguistics and artificial intelligence literature on spatial relations, but were unable to find descriptions for these relations which were adequate for the explanation of even the most simple cases of semantic distinction. Accordingly, we began a systematic inquiry into the ranges and boundaries of acceptable variation in spatial relations from which we could infer an algorithmic description.

Beginning with the simplest cases, we note that the example posed by Herskovits quickly leads to difficulties. In Figures 2 (a) and (b), circles A and B have a common angular deviation α from the horizontal axis, yet A seems describable as "directly to the right of" the square X, while B clearly is not prototypically "directly to the right of" X.

The problem shown in Figure 2 is not simply attributable to increasing distance. First, increasing distance should not be a factor in the angular deviation model, particularly at the relatively small change in distance in Figure 2. Second, a similar problem occurs in the case of decreasing distance. In Figure 3, why does rectangle C seem to be *below* square X while rectangle D is not prototypically *below* X? In both cases the angular deviation is identical, yet the prototypical spatial relation holds only for rectangle C.



(a) A is directly to the right of X, although offset by α.

(b) B is not directly to the right of X, although again offset by α.

Figure 2: With constant angular deviation, change in distance affects judgments about directional spatial relations.

(a) C is below X.

(b) D is not below X.

Figure 3: With constant angular deviation, change in distance affects judgments about directional spatial relations, yet the effect of the change in distance on the judgment is opposite that in Figure 2.

(a) E is left of Y.

(b) F is not left of Y.

Figure 4: Despite constant angular deviation, qualitative nature of the spatial relations changes markedly.

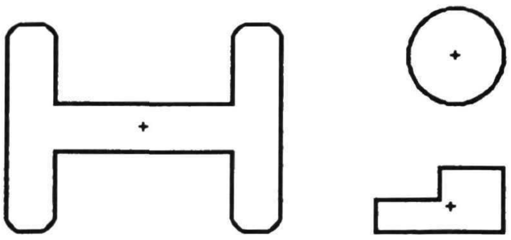


Figure 5: Centroids of various shapes.

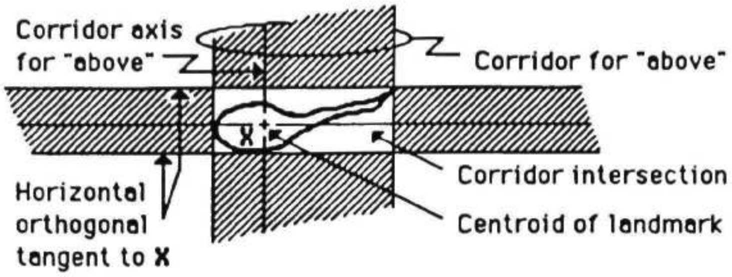


Figure 6: Boundaries and regions defined by a landmark shape.

Another problem situation for angular deviation is shown in Figure 4. In Figure 4 (a), the circle E is clearly directly to the left of the object Y. Yet in Figure 4 (b), with the same null angular deviation from the x-axis, circle F cannot reasonably be considered directly to the left of Y. One might say that circle E is *inside* object Y, but that is not a result predicted by constant angular deviation.

3 Fundamental Loci

To overcome these problems, we propose an algorithm which uses the following concepts:

1. Using terminology similar that of Langacker [1986], a spatial relation exists between a target and a landmark. The target is the object whose location an utterance seeks to declare. The landmark is a different object which is used to locate the target. Thus in the sentence "The circle is directly to the left of the square," the circle is the target and the square is the landmark.
2. The centroid of an object is the center of its area. That is, the centroid is like a flat object's "center of gravity," except that the object is assumed to have uniform density [Rosenfeld, 1976]. Thus the centroid of a circle is its center, the centroid of a symmetrical dogbone would be in the center of its shank, and the centroid of an asymmetrical shape is proportionately offset. See Figure 5.
3. The corridors of a landmark are the horizontal and vertical "shadows" of the landmark defined by its orthogonal tangents. As shown in Figure 6, the corridors of the landmark object X are indicated in gray. The corridor intersection is the rectangular region bounded by the landmark's orthogonal tangents. The centroid axis of a corridor is a line extending from the landmark's centroid in a direction parallel to the corridor.

With these concepts, we now propose an algorithm for deciding which term in the set *left of, right of, above,* and *below* best fits a given spatial relation (if at all).

```
IF the target does not overlap the landmark's corridor intersection
THEN  IF    (1) the centroid of the target is in a corridor,
      OR    (2) exactly one of the landmark's corridor
            contains any point of the target,
      OR    (3) the vector from the centroid of the landmark
            to the centroid of the target
            is closer to the centroid axis of a corridor
            than to the centroid axis of any other corridor,
      THEN  the relation between the target and the landmark
            corresponds to the direction of the corridor,
ELSE no particular spatial relation from this set is apparent.
```

The spatial relations algorithm is illustrated in Figure 7. Square A shows case (1); it is clearly to left of the landmark. Square B shows case (2); it is below the landmark. Square C shows case (3); it is more above than right of the landmark because the vector from the centroid of the landmark to the centroid of the target is closer to the *above* centroid axis than to the *right of* centroid axis.

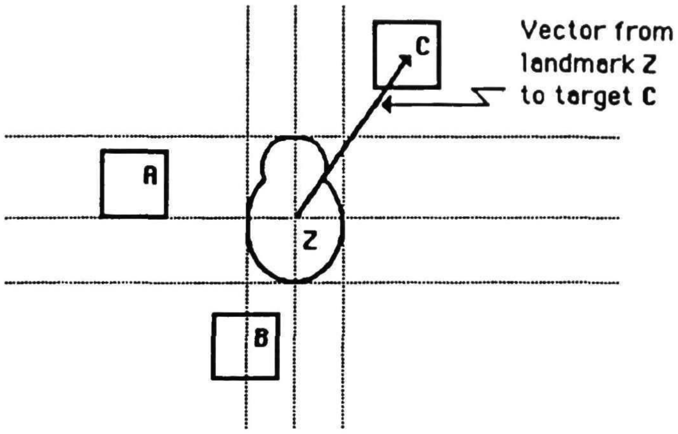
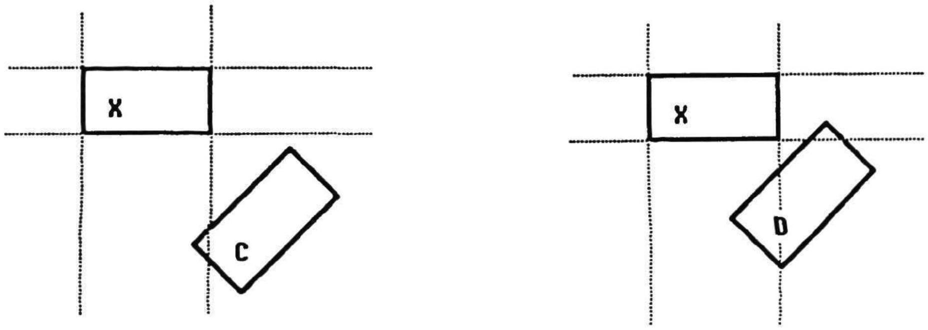


Figure 7: Cases of the spatial relations algorithm.

(a) The centroid of A is in the *right of* corridor of X.

(b) The centroid of A is not in the *right of* corridor of X.

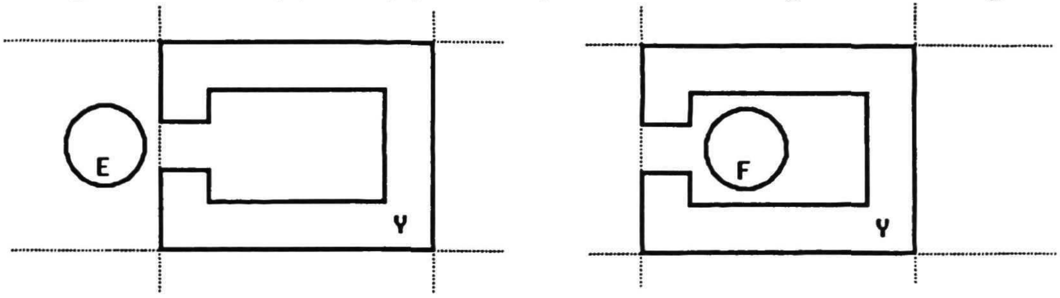
Figure 8: Cases (1) and (3) of the spatial relations algorithm distinguished.



(a) Only one corridor of X contains any points of C.

(b) More than one corridor of X contains points of D.

Figure 9: Cases (2) and (3) of the spatial relations algorithm distinguished.



(a) The centroid of E is in the *left of* corridor of Y.

(b) The centroid of F is in the corridor intersection of Y.

Figure 10: Applicability of the non-overlapping cases of the spatial relations algorithm.

Figures 8 through 10 show the algorithm applied to the examples of Figures 2 through 4. In Figure 8 (a), the centroid of target circle A is in the *right of* corridor of landmark square X, so case (1) applies. In Figure 8 (b), the centroid of target circle B is not in any corridor of landmark square X; at best, case (3) applies. In Figure 9 (a), exactly one corridor of landmark square X contains points of target rectangle C; this is thus case (2). Figure 9 (b) displays the corresponding case (3) situation. In Figure 10 (a) target circle E is clearly to the left of landmark object Y; E's centroid is in Y's *left of* corridor. In Figure 9 (b), however, circle F is in Y's corridor intersection, so none of cases (1), (2), or (3) applies.

This approach allows for variation from the prototypical loci, yet avoids the anomalous results of angular deviation as the standard. For example, as long as the centroid of the target is in a corridor of the landmark, the target's location can vary freely and yet the perceived spatial relationship will continue to hold. An important aspect of this approach is that it is not reducible to boolean evaluation of simple relational predicates. Rather, the approach seeks the best choice from the set of possible linguistically determined spatial relations based on the physically determined spatial relations in their semantic context.

4 Discussion

4.1 Theoretical Premises

Four theoretical or methodological premises have shaped both the criticisms and the alternatives we offer above.

4.1.1 Task Variation

First, the range of human performance in dealing with the language of spatial relations is broader than typically addressed. In particular, it is important to recognize that there are at least four different kinds of tasks involving spatial relations which subjects routinely perform. These tasks are illustrated in sentences (1) - (4). One can ask subjects to *verify* the truth of some sentence, as in (1). One can ask subjects to *choose* the most appropriate expression describing some state of affairs from a set of alternative, semantically "correct" expressions, as in (2). One can ask subjects to *describe*, either simply or in great detail, the relation holding between some target object and a landmark, as in (3). And, one can ask subjects to *manipulate* some part of the world in response to an utterance using a spatial relation expression, as in (4).

(1) The circle is above the square.

(2) Is the circle above or to the right of the square?

(3) Where is the circle located?

(4) Put the circle above the square.

Analysis of (i) seven protocols of two language teachers spontaneously teaching these spatial relations with geometric tiles and (ii) five protocols of subjects describing a videotape of moving circles and squares discloses no task outside this set.

Previous work typically focused almost exclusively (and implicitly) on the first kind of task, relying principally on introspective consideration of the grammaticality or felicity of sentences like (1). But while a subject will agree to the truth of (1) when shown Figure 8 (b), no subject ever places a circle in that location in response to the command in (4). A complete accounting, then, of the grammar of spatial relations requires that we account for this wider range of human behavior.

4.1.2 Lexical Alternatives

Second, the problem one faces in modeling spatial relations is not one simply of specifying a precise meaning or rules of use for individual lexical expressions in isolation from each other. Rather, within a given semantic domain, such as space, the speaker makes choices among acceptable linguistic expressions, identifying the most appropriate expression for the given context. So, for example, Figure 8 (b) can be described either by (6) or by (7),

(6) The circle is to the right of the square.

(7) The circle is above the square.

for each correctly describes the state of affairs represented by Figure 8 (b), and subjects asked to verify the truth of either sentence alone will accept each as true. But of the two, (7) is recognizable as more appropriate. The use of linguistic expressions, then, is not wholly determined by the meaning of a given expression alone but in concert with semantically associated items in the same domain.

4.1.3 Semantics or Pragmatics

Third, previous approaches to spatial relations tend to conflate semantic facts concerning the intrinsic meaning of expressions with pragmatic issues concerning their use. In the approach taken here, the semantics of an expression represents the full range of cases for which the expression is true. The inherent semantic meaning of a lexical expression for a spatial relation like *left* or *above* is represented by the region whose boundaries are defined by the cases for which the expression is true versus those for which the expression is false. Methodologically, it is the verification task that permits one to discover exactly what that region is. This approach motivates the basic notions used in the algorithm.

The pragmatics of linguistic expressions represents the principles governing the choice among alternative true expressions. In our approach, one selects the best-fitting expression arising from the domain. In pragmatics, the other three kinds of linguistic task described above are also important. Choice tasks help determine the boundaries of alternatives in conflict. Manipulation tasks help determine optimal choices. Description tasks help determine the variables relevant to pragmatic principles of use.

Overall, we have found that a strict separation of semantic and pragmatic issues allows one to account for the wide range of cases represented by the four task types. Further, there appears to be no need to formulate prototype meanings for any expression, for the prototype effects fall out from this treatment.

4.2 Task Variation and the Proposed Algorithm

If variation in the use and comprehension of linguistic expression of spatial relations depends in part on task, then task serves as one predictor for acceptance of deviation from any prototypical locus. The algorithm presented above addresses some of the variability in use of expressions due to task variation.

For manipulation tasks, the locus of acceptable points for a spatial relation corresponds to a special case of case (1) in the algorithm. This reflects the case, for example, in which one tells the subject "Put the circle to the left of square." We expect that the subject will typically put the target relative to the landmark so that the centroid of the target is on the centroid axis of the landmark's left corridor, and so that the target is near the landmark without overlapping. The meaning of *near* is the subject of related work [cf. Denofsky, 1976].

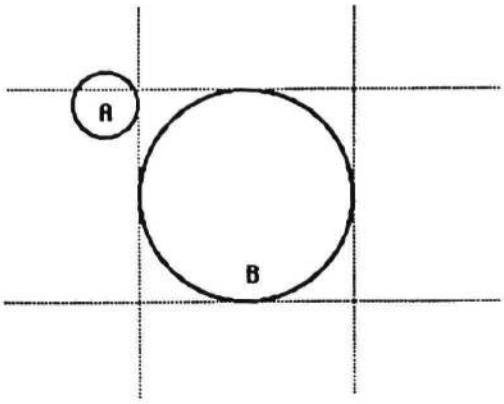


Figure 11: Spatial relations can be non-commutative.

For the description and choice tasks, a subject spontaneously uses the best term to describe a spatial relation between a target and landmark. This task corresponds to case (1) of the algorithm. If the centroid of the target is anywhere in the left corridor of the landmark, then we expect a subject to describe the relation between the objects as "The circle is to the left of the square." (This assumes that the objects don't overlap.) Anywhere else, the subject chooses the closest vector.

For the verification task, the subject chooses a truth value for a statement about the spatial relation between objects. This task corresponds to cases (2) and (3) of the algorithm. Subjects will find case (2) clear, and will find case (3) less clear as the vector from the landmark to the target becomes equidistant from two neighboring centroid axes of the landmark. In Figure 7, with respect to the landmark object Z, square A is in a case (1) relation, square B is in a case (2) relation, and square C is in a case (3) relation. Note that the location of square A corresponds to a prototypically correct response to the command "Put the square to the left of Z." Likewise, the locations of squares B and C will typically elicit an affirmative response to the the question "Is the square above (or below) Z?"

Note that these tasks require an explicit identification of target and landmark. If these roles are reversed for a given display, the perceived spatial relation between objects may not be maintained. This lack of commutativity is reflected in the algorithm. For example, in Figure 11, a subject would probably describe circle A as being to the left of circle B, but would not describe circle B as being to the right of A.

4.3 Function and Structure

The spatial relation algorithm presented above, on its terms, does not apply to many possible cases, especially where the target lies wholly or in part in the landmark's corridor intersection. Evidence we have collected suggests that many of these cases can be resolved by reference to world knowledge of function and structure. By function, we mean uses of objects, such as containment. By structure, we mean the constituent parts of an object which make up its shape (and the relation of these parts to each other, of course).

Just as the task in which objects are involved influences the nature of their perceived spatial relations, so too may the function of objects affect spatial relations. In Figure 4 (b), the target circle would be best described as *in* the landmark object. However, simple geometric definitions of *in* and *out* are not adequate for domains which purport to model real-world semantics. While the landmark object in Figure 4 (b) has no apparent semantics aside from its polygonal structure (and perhaps the quality of being a reversed letter "C"), we believe that the real-world semantics of objects such as those in Figure 12 would overwhelm the naive geometrical approach. At this point, then, we represent the function of objects by explicit reference to their functions. For example, a thing is or is not a *container*.

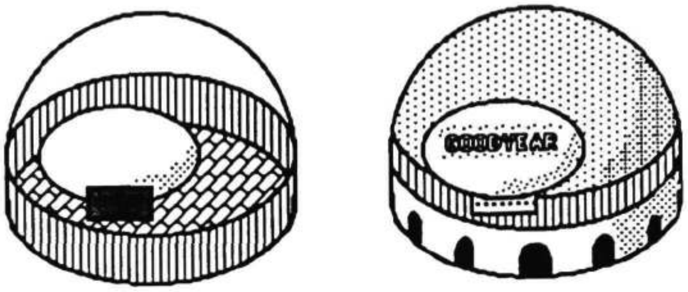
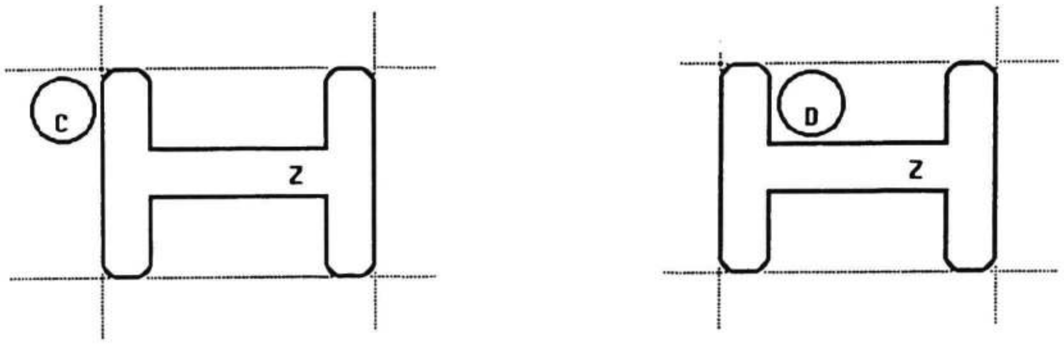


Figure 12: The domain semantics of shapes affect their spatial relations. Thus identical shapes (in outline) in the same physical relation in two-dimensional representation will create different apparent linguistic relations depending on the domain information imparted by their internal representations.



(a) Circle C is *left of* shape Z.

(b) Circle D is not *left of* shape Z, but rather "above the shank just to the right of the left knob of Z."

Figure 13: Local spatial relations used where the target overlaps the corridor intersection of the landmark.

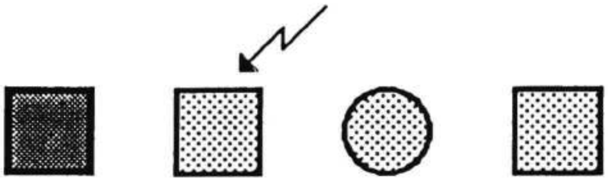


Figure 14: Spatial relation disambiguated by combining shapes in an implied global structure. The arrow indicates for the shape to which the experimenter pointed; the arrow was not present in the figure actually presented to subjects.

With respect to structure, we noticed that where subjects had difficulty describing a relation (typically because the target lay within the landmark's corridor intersection), the subject would resolve the problem by identifying a part or parts of the landmark and then using the usual spatial relations expressions relative to those parts. A typical example is shown in Figure 13. In Figure 13 (a) the target circle C is unambiguously *left of* the landmark object Z, and subjects will give the expected response. But when faced with Figure 13 (b), subjects will produce a description like "The circle is above the shank just to the right of the left knob." That is, the subject has broken down the landmark into parts from which the normal discourse of spatial relations can be constructed. Interestingly, we also observed a converse effect in which ambiguity among spatial referents was avoided by ideating a structure comprised of the referents and then constructing an utterance based on a spatial relation involving the structure. For example, when Figure 14 was presented and the indicated object pointed out, subjects typically identified the object as "the second square from the left," impliedly creating a horizontal structure encompassing all four objects. Subjects typically did not identify the object as "the square to the left of the circle."

4.4 Future Work

Our future work in this area involves, among other things, systematic experimental verification of the spatial relations algorithm. We plan to present subjects with a variety of tasks and relations from which we can extract data on manipulation, description, choice, and verification, with reaction times where appropriate. This work will be conducted on a Symbolics 3645 Lisp machine using testing software specially developed for these series of experiments.

Other work will include continued research into the meaning of *near*. We expect to apply similar techniques to this term. We are also continuing to analyze tasking in the discourse of spatial relations. The algorithm that we have presented in this paper, while inherently accommodating the semantics if *left of*, *right of*, *above*, and *below*, nevertheless lacks explicit procedures for generating the pragmatic aspects of task, function, and structure that we have discussed. Our future work will focus on that elaboration and its empirical validation.

5 References

Denofsky, Murray Elias (1976). *How Near is Near?*, AI Memo No. 344, Massachusetts Institute of Technology Artificial Intelligence Laboratory, February 1976

Ehrich, V. (1985) The Linguistics and Psycholinguistics of Secondary Spatial Deixis. In G.A.J. Hopenbrouwers, P.A.M. Seuren, and A.J.M.M. Weijters, *Meaning and the Lexicon*. Foris Publications: Dordrecht, Holland.

Hayes, P. (1985b). Naive Physics I: Ontology for liquids. In J.R. Hobbs and R.C. Moore (Eds.), *Formal Theories of the Commonsense World*. Norwood, NJ: Ablex Publishing Corp.

Herskovits, A. (1985). Semantics and pragmatics of locative expressions. *Cognitive Science, 9*, 341-378.

Langacker, R. W. (1986). An Introduction to Cognitive Grammar, *Cognitive Science 10*, 1-40.

Rosenfeld, A., & Kak, A. C. (1976). *Digital picture processing*. New York: Academic Press.

Talmy, L. (1983). How language structures space. In H. Pick and L. Acredolo (Eds.) *Spatial Orientation: Theory, Research, and Application*. New York: Plenum Press.