

# UC Santa Barbara

## NCGIA Technical Reports

### Title

Spatial Relations Representation and Locative Phase Generation in a Map Context (89-14)

### Permalink

<https://escholarship.org/uc/item/48z974t6>

### Author

Haller, Susan M.

### Publication Date

1989-12-01

**National Center for Geographic Information and Analysis**

**Spatial Relations Representation  
and Locative Phase Generation  
in a Map Context**

by

**Susan M. Haller**

National Center for Geographic Information & Analysis  
Department of Computer Science  
State University of New York at Buffalo

National Center for Geographic Information & Analysis/NCGIA  
**Technical Paper 89-14**  
**December, 1989**

(Re/ 12)

## **Acknowledgements**

This paper represent parts of Research Initiative #2, "Languages of Spatial Relations", of the National Center for Geographic Information and Analysis, and was supported by a grant from the National Science Foundation (SES8810917). Support by NSF is gratefully acknowledged.

# Spatial Relations Representation and Locative Phrase Generation in a Map Context\*

Susan M. Haller  
Department of Computer Science  
State University of New York at Buffalo  
226 Bell Hall  
Buffalo, NY 14260  
(716) 6363182.

## Abstract

We are concerned with generating natural language locative phrases with respect to a map in response to "Where is "type queries by a user. We restrict our problem by assuming that the correct ground (reference point) to use for locating a figure (place being located), is a stationary landmark of enough size and importance that is visible on the map. Hence, our problem involves choosing a suitable landmark for the ground, building a knowledge representation to express the relationship of the figure to the ground, and generating natural language to express that relationship. Our work has resulted in an extension to a generation grammar for a multi-modality interface system that selects a reference point and expresses the location of an object with respect to the selected reference point. In doing so, the system attends to what is prominent on the map, and what has been used recently to orient the addressee. The development of our extension has clarified several research issues related to generating natural language consistent with, and appropriate to visual information displayed on map.

Topic Areas: *knowledge representation, spatial relations, semantic networks, natural language generation, locative phrases, map systems*

Domain area: *Geographic information systems, map systems.*

Language/Tool: *SNePS (Semantic Network Processing System), CUBRICON (CUBRc Intelligent CONversationalist)*

---

\* The research discussed in this report represents part of Research Initiative #2, "Languages of Spatial Relations," of the National Center for Geographic Information and Analysis, and was supported in part by a grant from the National Science Foundation (SES8810917).

# 1 Introduction

## 1.1 An Overview of CUBRICON

This work was conducted by extending an existing generation grammar in a multi-modality interface system called CUBRICON (the CUBRc Intelligent CONversationalist) [3]. This system accepts natural language input that can include pointing gestures via a mouse to a map display. The output of the system is natural language, most often in conjunction with pointing and blinking gestures on the map.

The CUBRICON map system has two visual and hence, representational planes. A map of East/West Germany with visual representations of cities, towns, rivers, highways, lakes, power lines, etc. is available at various levels of refinement obtained by zooming in on sub-regions of the map. Names and other geographic information about these entities are also present at this underlying level.

The so-called icon level of the map overlays the map of East/West Germany. Entities represented visually at this level by icons are military and military related: air bases, factories, fuel storage facilities, missile sites and various military deployments. Information associating these various entities to appropriate icons, and other military and geographic information on the entities is included in a knowledge base represented in the Semantic Network Processing System (SNePS)[5].

There are three main maps that the user can ask to have displayed. The East/West Germany map displays the entire region that the CUBRICON map system is responsible for. Two additional maps of the Fulda Gap and Leipzig regions display strategic sub-regions of the East/West Germany region.

The natural language understanding and generation components are implemented in a Generalized ATN grammar [4] and are used for analyzing input sentences and generating English responses. In response to input requests for various displays, or queries for locations of entities at the icon level, the system finds the appropriate semantic node in the knowledge base that represents the information, and expresses it via natural language and associated pointing/blinking gestures if appropriate.

## 1.2 Locative Phrases

Herskovits defines a *locative expression* as any spatial expression involving a preposition, its object, and whatever the prepositional phrase modifies (noun, clause, etc.) [2]. The prototypical use of the expression is to inform the addressee of the location of the modified noun or clause. Herskovits refers to this located object as the *figure*. The object of the preposition, called the *ground*, is the reference object used to specify a spatial relationship between itself and the figure. It is assumed that the location of the ground object is either familiar to, or easily discovered by the addressee. In the locative expression,

*the house on the hill*

*house* is the figure object whose location is constrained by the ground object *hill* through the spatial relationship *on*.

We note that besides locating figure objects, locative expressions serve as descriptors as in *the book on the shelf*. There is also a broad range of metaphorical use, such as *Tom is in trouble*. However, our concern with these expressions, here, is limited to their use in locating figure objects, and we simplify our problem further with the constraint that all such utterances occur in reference to a two-dimensional map.

## 1.3 Generating Locative Phrases in Response to "Where is" Queries

In the CUBRICON system, the user can ask for information on the location of any entity in the knowledge base that has a name. Such entities are air bases and factories that are located throughout the East/West Germany region, and since they are of primary interest, their iconic representations are almost always visible on the map.

Currently, when the user asks where such an entity is (via a "Where is" query), if the site is viewable or can be made viewable on the map, the system responds by blinking, highlighting and labeling its iconic representation. In addition, a canned phrase is printed out:

*The < name and class of object > is located here.*

If the site is outside the region represented by the map system, a message is printed to that effect, and the node corresponding to the site's location in longitude and latitude is handed to the natural language generator:

*Its location is < latitude and longitude of site >*

The goal of this project is to generate a felicitous locative phrase that expresses a site's location relative to some landmark viewable on the map. We attempt to express the whereabouts of the figure site in terms of the map in those cases where it can be seen, and in those cases where it cannot be seen. Generating such a locative phrase involves three distinct steps:

1. Choosing an appropriate landmark as the ground object
2. Determining a spatial relationship between the figure and ground object
3. Generating natural language to express the spatial relationship

We will present representations and algorithms as they relate to the discussion of each of these steps.

## 2 Step One: Choosing a Ground Object

As previously mentioned, we assume that appropriate candidate ground objects are geographic entities of the East/West Germany region of enough size and importance so as to be viewable on the map and familiar to the user. With this heuristic in mind, seven major cities located throughout the East/West Germany region were selected and represented in the SNePS knowledge base.

Since SNePS is a propositional semantic network, information is represented by nodes and arcs. In Figure 1, square nodes with no outgoing arcs denote constant individuals of thought. The one exception is nodes at the ends of *lex* arcs, these nodes represent interfaces to the external world, here, as words. Circular nodes correspond to propositional nodes that represent information about individuals and concepts (also represented by circular nodes). With respect to Figure 1, individual cities all belong to the class "city" which is associated with a small, blue point-type icon (Figure 1). In Figure 2, each individual city is associated with its name and location in longitude and latitude.

In addition to cities, the East/West Germany border was added to the knowledge base for use as a possible reference point. It was chosen for its political relevance and because it is visually prominent on the map of East/West Germany and on the maps of the two sub-regions. Usually in view, this border serves as a good landmark when no cities can be seen. It is represented as a path-type entity in the knowledge base (Figure 4) where a path value is a series of points, each with location-type values of their own (Figure 6). The border is a member of the class "border," (Figure 3), the series of points are all members of the class "border point," (Figure 5), which is also associated with the same small, blue point-type icon. The border has one additional property associated with it, the division property. This is an internal value used for determining the direction of a figure from the border as a ground and will be included in the discussion of determining spatial relationships.

The following heuristics are employed to select a ground object:

- The ground object selected must be visible without expanding an existing window or creating a new one.
- Landmarks previously used as ground objects are preferable to new ones if they are reasonably near the figure object.
- If a new landmark must be used as ground and there is more than one, select the landmark nearest the figure object.

City Class

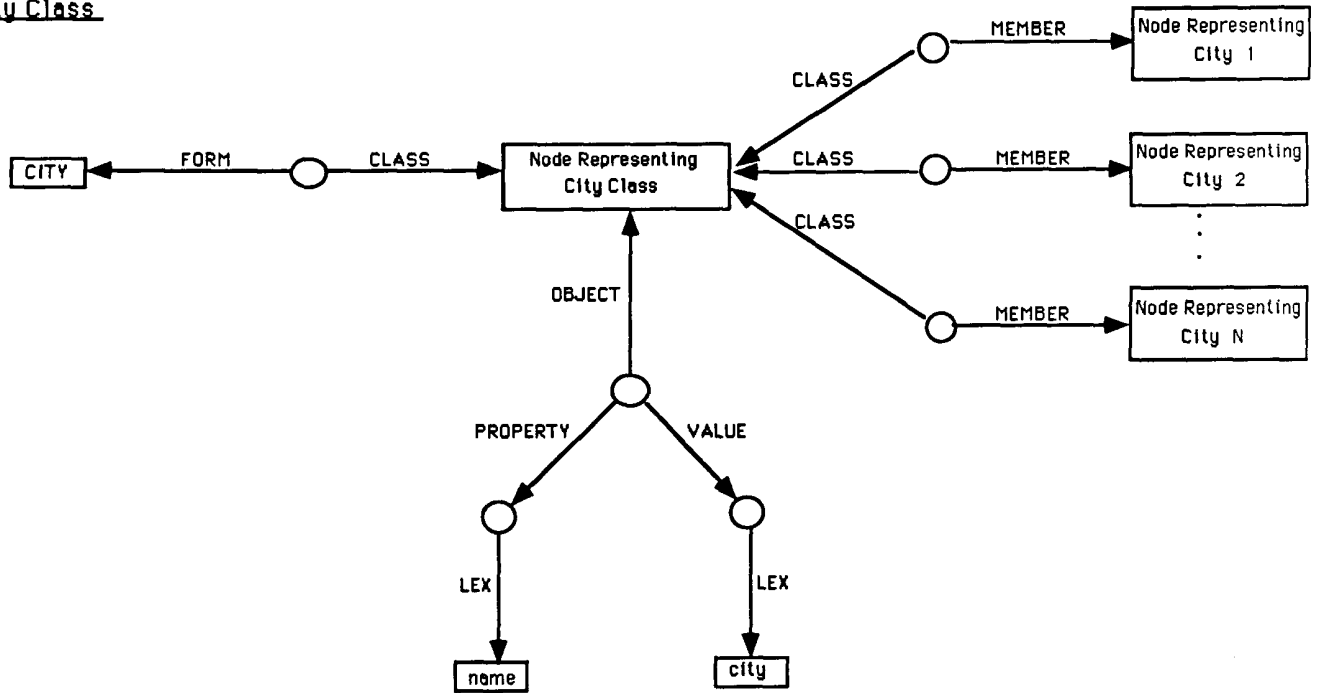


Figure 1: Representation of a City Class

Sample City

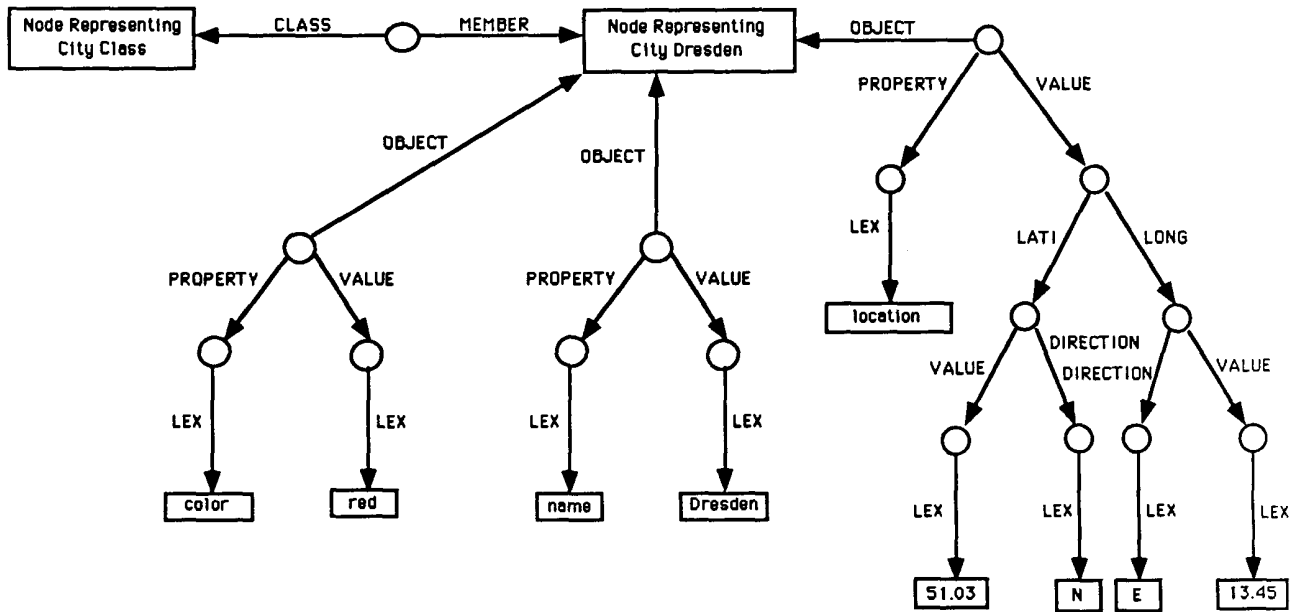


Figure 2: Representation of a Sample City

Border Class

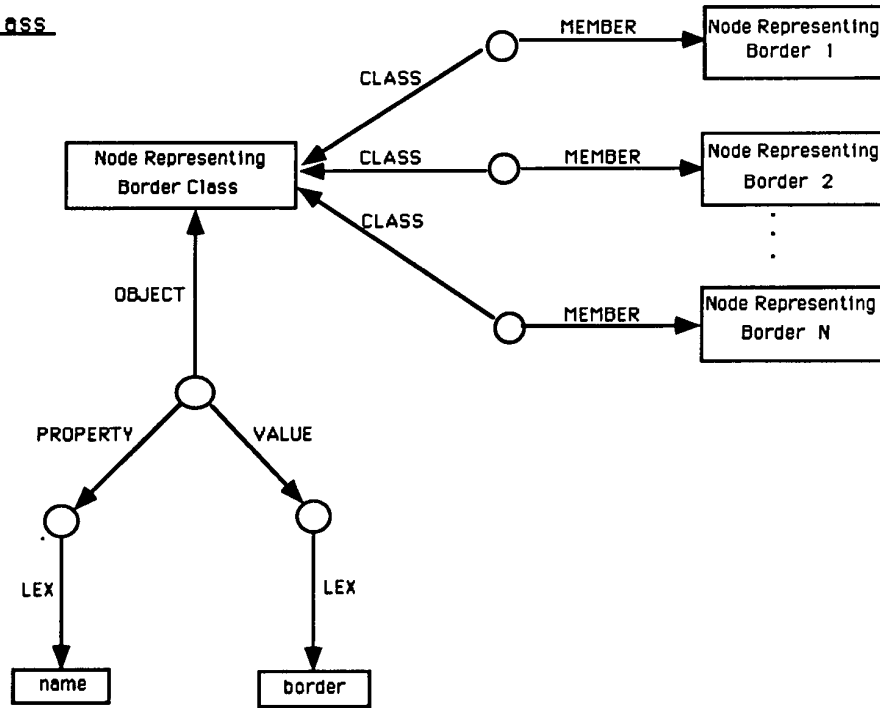


Figure 3: Representation of the Border Class

Sample Border

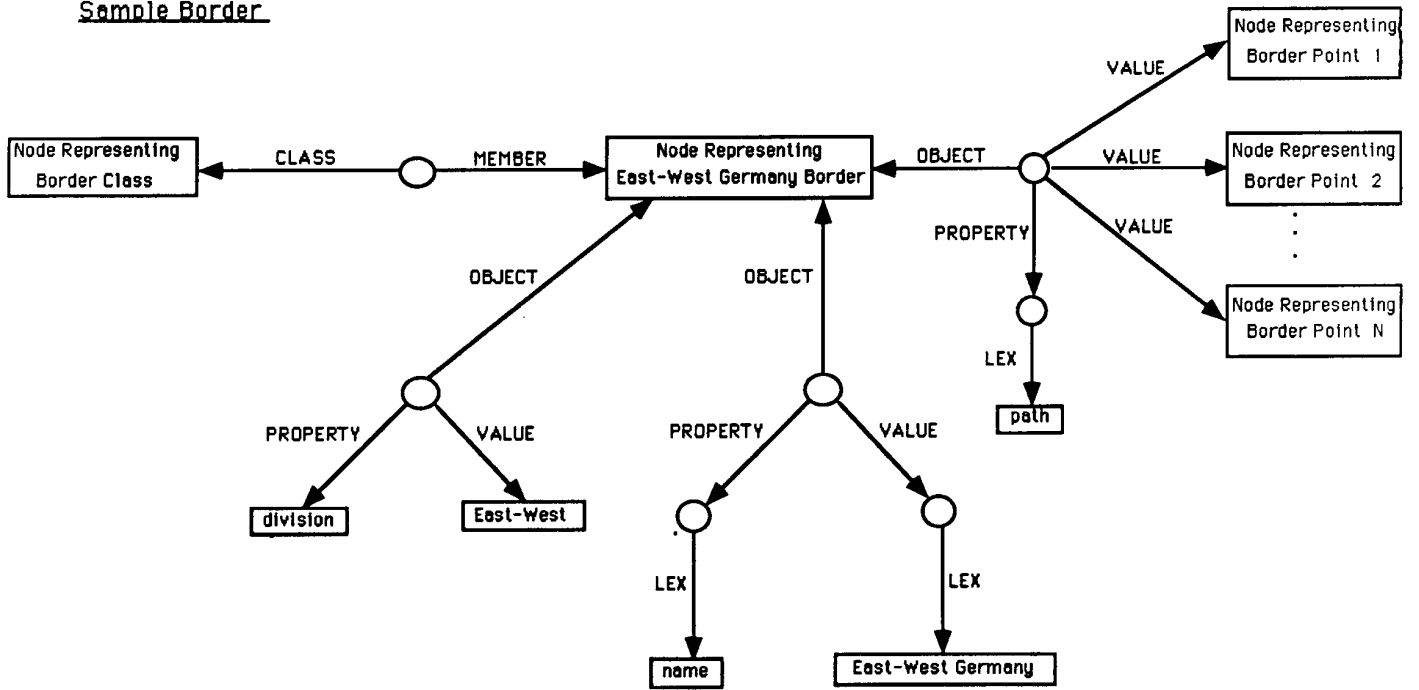


Figure 4: Representation of a Sample Border



Border Point Class

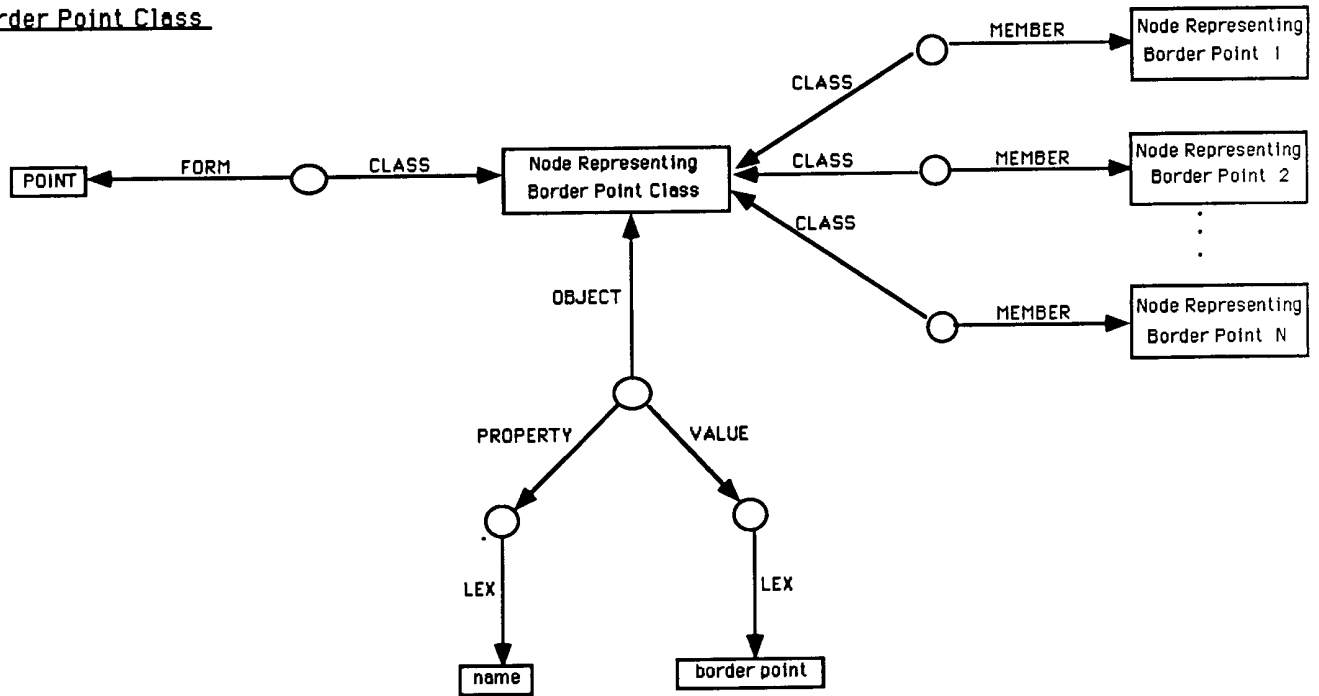


Figure 5: Representation of the Border Point Class

Sample Border Point

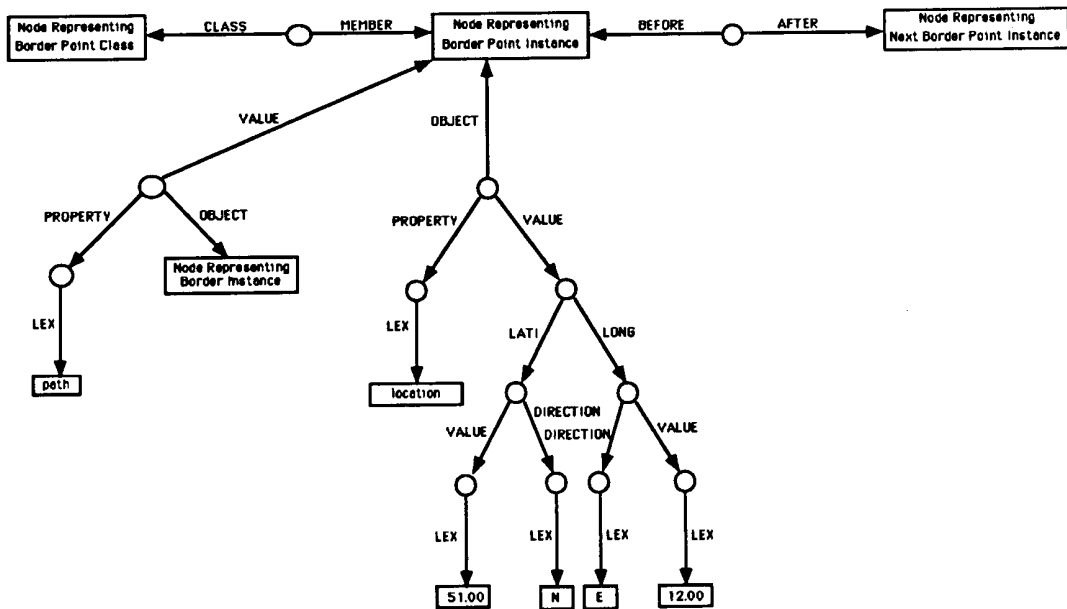


Figure 6: Representation of a Sample Border Point

The CUBRICON system is capable of expanding or changing maps when a request is made to the icon level of the system for an entity that is viewable, but perhaps not currently displayed. It is our opinion that reference points are objects that are available at no effort to help orient the user about the whereabouts of the figure. Hence, the algorithm we use carefully checks to make sure a candidate ground object is viewable within the boundaries of the current window that contains the figure *before* the iconic representation of the ground object is added.

The natural language component of the CUBRICON system maintains a list of *discourse items* on a *focus list* [6]. The list is primarily used to decode and encode pronoun referents, however we put it to one further use. Landmarks previously used as ground objects are viewed as particularly salient not only for being a part of the addressee's mental discourse model, but also for being previously cast in the role of a ground object. If it has recently been indicated to the listener that *the Lindsey air base is west of Frankfurt*, they will be particularly receptive to the use of this ground again, if appropriate, and if it is used within the next two or three exchanges:

User: Where is the Lindsey air base?  
System: Its location is west of Frankfurt.  
User: Where is the Nurnberg air base?  
System: Its location is southwest of Frankfurt.

Therefore, when a landmark is used as a ground object, it is added to the focus list with sufficiently low priority so as to be removed from the focus list within a few exchanges. However, for a short time the landmark is available for use, and if it is in the same window as the current figure object, and if it is "near"<sup>1</sup> the figure, it will be selected once again as the ground object.

If there are no appropriate landmarks in focus the third heuristic embodies the bottom line choice strategy among available landmarks. Currently, all landmarks in the window with the figure are viewed as equally good candidates for the ground, so unless one of them has been recently used as a reference point, we use the closest landmark to the figure.

As previously noted, some figure objects are not viewable. In this case, the algorithm attempts to find a region (the East/West Germany, Fulda Gap or the Leipzig) that has been previously displayed and hence, is on the focus list. Unlike other landmarks, regions added to the focus list remain there for a long time since they are added in the subject syntactic-semantic category as a direct result of the "display" region request by the user. We consider such areas to be valid landmarks since they can, and have been asked for by name and axe therefore familiar to the user. Hence, if the algorithm succeeds, the region itself is used as the ground to orient the user to the location of a figure that cannot be viewed.

We note that the algorithm can fail to find a suitable landmark in the following situations:

- The figure is outside the area represented by the map system or is for some reason not displayed and any region displayed previously is no longer on the focus list.
- The figure is visible in a presentation window that contains no landmarks, most likely because zooming in on the figure has occurred.

Under these circumstances, the system defaults to reporting the figure object's latitude and longitude. We note that in the second case, the system could locate the object with respect to the window itself (i.e. upper right hand corner, left side, etc.), however this was judged to be inappropriate since greater geographic detail is available in enlarged sub-regions, and the appropriate landmarks change from cities, to towns, to roads and so on. A better solution is to use landmarks appropriate to scale, but this has not been addressed yet, so the case is purposely left open-ended.

To conclude and summarize this section, we give the algorithm used to choose a ground object from candidate landmarks. This algorithm is implemented in the top-level function called by the generation grammar to find a landmark and build and return a relative relation.

### **Choose-ground-object:**

IF the figure is not within the region represented by the map  
OR not otherwise displayed THEN

IF there is a named region that has been displayed recently THEN use it as the ground  
ELSE (\* any region that could be used has not been used in a while\*) return nil as the ground (\* use latitude and longitude\*)

---

<sup>1</sup> "Near" is currently calculated as within a half a window width, a perceptually valid metric is needed, however.  
Choose-ground-object:

ELSE IF the figure is being displayed in a window THEN

IF there is a landmark in focus that is in the same window and "near" the figure THEN use the landmark in focus as the ground

ELSE IF there are landmarks in the same window as the figure THEN use the landmark closest to the figure as the ground

ELSE (*\* there are no landmarks in the window with the figure\**) return nil as the ground (*\* use latitude and longitude \**)

### 3 Step Two: Determining and Building the Spatial Relationship

This problem is greatly simplified since we restrict our references to a two dimensional map. An additional constraint, currently, is the scale of the map. Only seven cities and the East/West Germany border are available as landmarks, and if the figure is enlarged sufficiently, all these landmarks may disappear and our current algorithm fails. Hence, we assume that if the algorithm finds a ground object, it is usually on a large scale map.<sup>2</sup>

In the context of a large scale map, the spatial relations we judge to be suitable are *east of, west of, north of, south of*. Since distance between figure and ground is desired, the information assumes a vector-like quality, and greater resolution in terms of the direction seemed appropriate, therefore *southwest of, southeast of, northwest of and northeast of* were added to the set of spatial relations. We note that these relations are special in that they are not from the set of English prepositions usually used to denote spatial relations such as *on, in, at*, etc., relations such as *northwest* are exceptional in that other relations do not express relations in terms of two dimensions.

As previously mentioned, the figure object is always a stationary military site with a proper name. In the knowledge base, these entities are associated with a location expressed in longitude and latitude in the same way as cities (see representation of a sample city, Figure 2). Since the location of the figure is always a coordinate pair, only the locational property of the ground needs to be considered in determining direction and distance to the figure. As discussed under ground selection, the ground object is either a city, a border or a region, and to express their location, these entities possess "location," "path" and "boundary" properties respectively. Hence, we use different procedures to compute a distance and direction from each of these types of ground objects to a figure.

#### 3.1 Computing Direction and Distance from a Ground Object with a Location Property

We took a straightforward approach to this case. The figure and the ground objects both possess location properties that are point coordinates expressed as a latitude and a longitude. The direction from the ground to the figure is calculated by placing the ground at the origin of a coordinate plane and calculating the angle  $\theta$  to the figure. The possible values of  $\theta$  (0 360 degrees) are broken up into octants, where each octant specifies one of the eight directions from ground to figure (Figure 7).

The calculation of direction in this way does not appear to always produce satisfactory results. The octants corresponding to northwest, northeast, southwest and southeast appear to take up more visual angular space than those corresponding to north, south, east, and west. For example, we seem to perceive the figure as being northeast or southeast of the ground more readily than east unless the figure is "almost" directly east of the ground. Determination of the perceptually correct division of this angular space is a topic for further work.

The distance from ground to figure is calculated by first converting the change in latitude and longitude to miles and using these distances as the legs of a right triangle to determine the length of the hypotenuse. We drop distance as part of the spatial relation entirely if the figure is perceived as being so close to the ground as to appear next to it. Here, just the direction from ground to figure is reported. The relative rounding or elimination of distance is expressed succinctly in the following algorithm implemented in the code as the function *refineorientation*.

##### **Refineorientation:**

IF the distance is less than 0.04 of the window width, THEN (*\* the figure appears next to the ground \**) drop distance as part of the spatial relation reported

ELSE IF the distance is greater than 0.10 of the window width, THEN (*\* the distance is a large portion of the display, which represents probably at least 50 miles \**) round to the nearest 10 miles

ELSE (*\* distance is > 0.04 but < 0.10 of the window width\**) round to the nearest mile

---

<sup>2</sup> Of course, "large scale is relative, here we take it to mean a map where the width of the map represents anywhere from 50 to 250 miles.

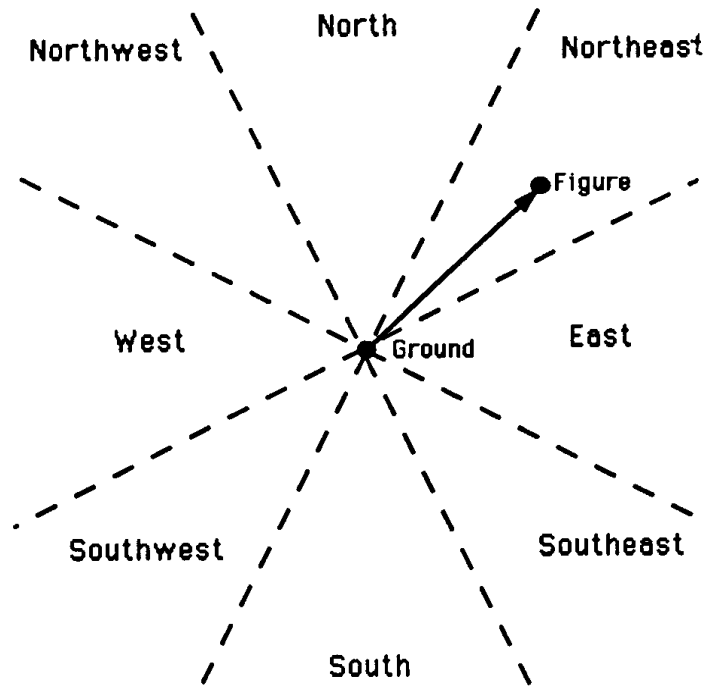
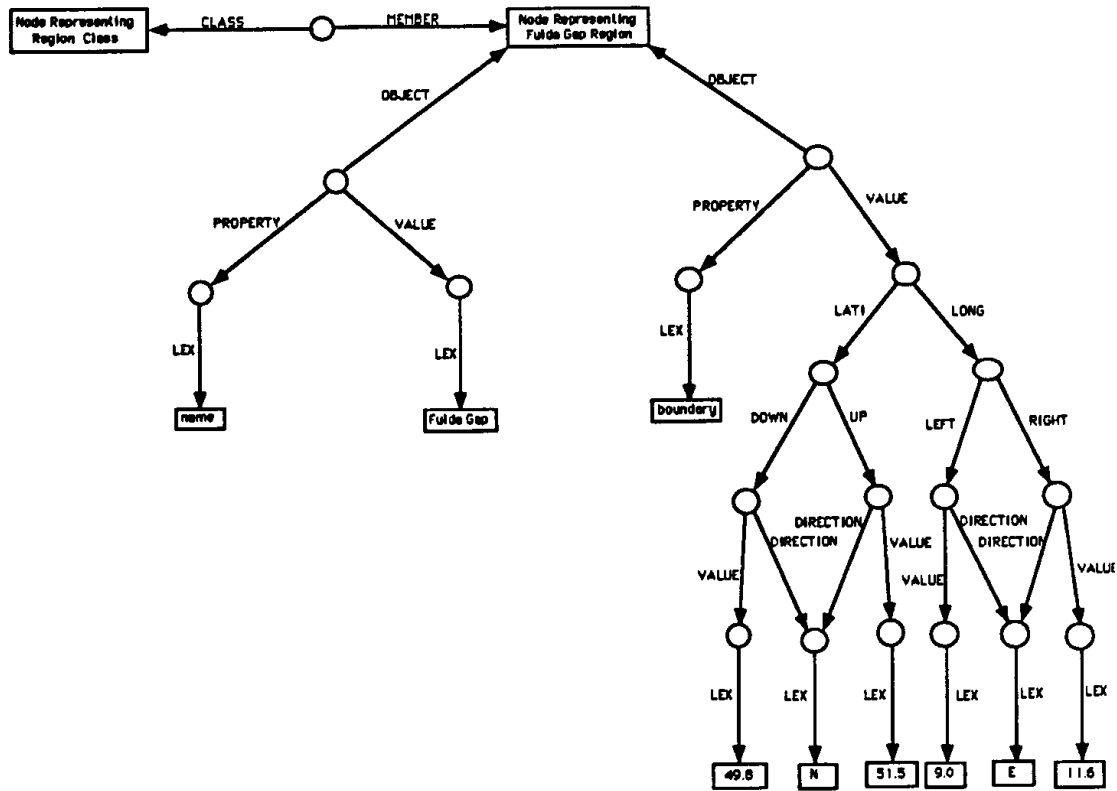


Figure 7: The Division of Angular Space for Direction Determination from a Location-type Ground

Sample Region



**Figure 8: Representation of the Fulda Gap Region**

We note that this heuristic assumes that the window width represents a scale of between fifty and five hundred miles. The choice of what to round to is also a function of the window width, however it should always be to the nearest mile or to some number that is has a "usual" factor such as five or ten (i.e. we don't report a result with rounding to the nearest 16 miles).

**3.2 Computing Direction and Distance from a Ground Object with a Boundary Property**

Recall that in those cases where a figure is not viewable, a named region is used as the ground object to orient the user about the whereabouts of the figure with respect to the viewable domain. This calculation was simple and produced satisfying results. Note that the ground is given as a rectangular region with a boundary specified as two longitude lines and two latitude lines. The representation of the Fulda Gap region is given as an example in Figure 8.

To calculate the direction to the figure from a rectangular ground, the boundaries of the ground were extended to create nine areas (the named region at the center and eight unbounded areas surrounding it, see Figure 9). The surrounding area in which the figure is located determines the direction from the region.

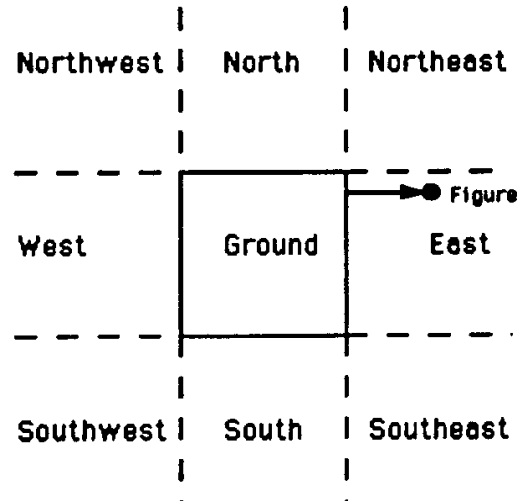
The distance to the boundary of the region is reported rounded to the nearest ten miles. We view this as appropriate since the purpose of the expression eventually generated from the spatial relation is to orient the user and not provide exact locational information. If the figure is not viewable, precise distances (for example, six miles vs. seven miles) are meaningless because they cannot be seen (and used) to develop a sense of the map scale. In addition, figure objects are never outside the region of responsibility by more than fifty miles. With these considerations in mind, we selected 10 miles as a cognitively valid rounding value.

**3.3 Computing Direction and Distance from a Ground Object with a Path Property**

Direction and distance determination from a path-type entity to a point-type figure object proved to be a difficult orientation task. Figures 4 and 5 show the border has a path property, the value of which, consists of a series of points that determine a linear

piecewise representation of the border as a path. The following observations are made regarding determination of direction from a border embodied as a path to a point-type figure:

- Most international borders, because of their political relevance and the vagueness with which they are perceived, predispose the user to expect one of only two directions.
- Some snapshots of highly irregular borders visually suggest very different directions than the ones expected.



**Figure 9: The Division of Space for Direction Determination from a Boundary-type Ground**

International boundaries are cognitive entities most people have some notion of independent of a map. Depending on the area of the world the user has knowledge of, some boundaries conjure up stronger images than others. For example, it is our opinion that most people with any geographic knowledge would say that the US/Canada border divides the continent into north and south, although clearly it divides some subregions into east and west. Hence, borders impose a certain division expectation which we represent as a proposition internal to the system, the division property (Figure 4).

In Figure 4, this information gives the system access to the notion that the user is expecting the figure to be located at a direction of east or west from the East/West Germany border. We attempted to build into the system some ability to resolve the direction of the figure from the border into east or west unless the visual evidence for north or south is very strong. For example, when the Fulda Gap region is in view (Figure 10), the Erfurt air base appears to be both east and north of the East/West Germany border. The airbase lies north of the closest point on the border, however, as shown in Figure 10, we would like the system to use a reference point that is east of Erfurt airbase, since an east/west orientation is preferred over north/south.

When the same border and air base are viewed in the Leipzig region (Figure 11), the air base is best described as north of the border, since the border points in view clearly divide the region into north or south. Here, the system should override its preference for east or west and respond with north as the direction from the border to the Erfurt air base.

When and how to resolve orientation to a border in this way proved to be a much difficult problem than anticipated. Algorithm choices that worked well in one or two situations produced poor results in others. As a result, we abandoned the attempt to use a direction preference and selected a simple algorithm that produced the most pleasing results overall. We find all the points on the border that are viewable in the current window and select the one that is closest to the figure object as the ground reference point. The distance and direction to the border is computed from this reference point to the figure. With respect to Figure 11, the system would express the location of the Erfurt air base as:

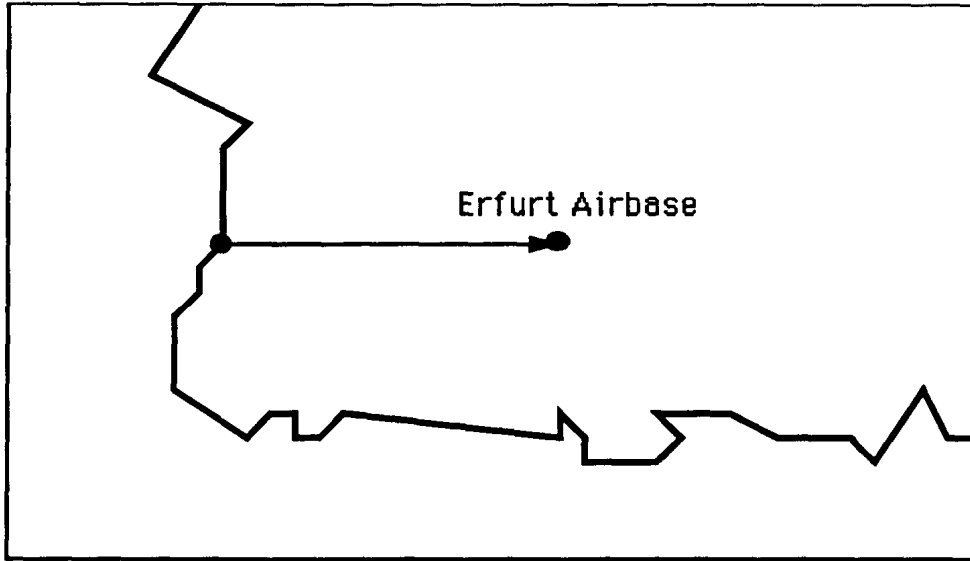


Figure 10: The East-West Germany Border and the Erfurt air base roughly as viewed in the Fulda Gap region.

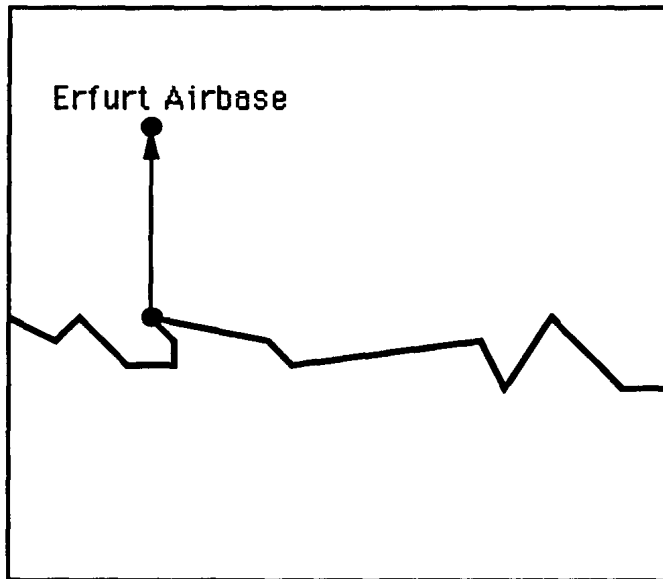


Figure 11: The East-West Germany Border and the Erfurt air base roughly as viewed in the Leipzig region.

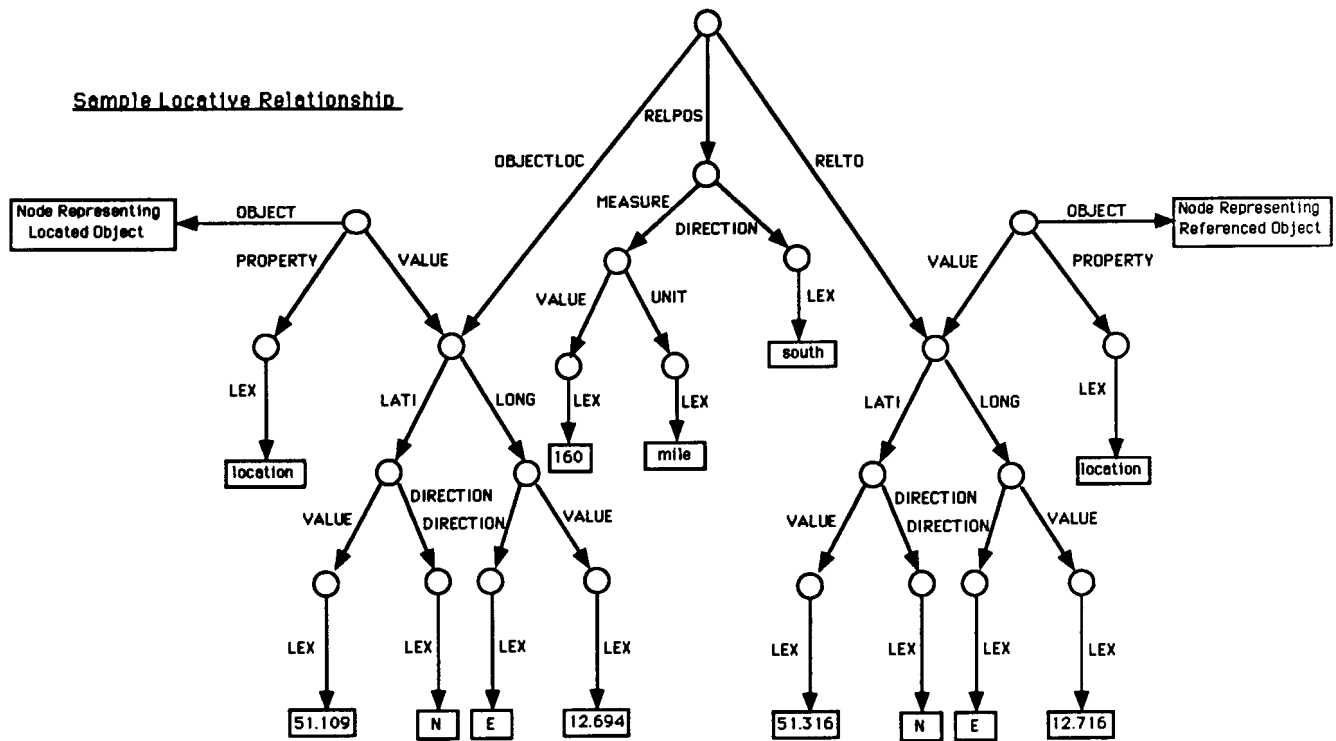


Figure 12: Representation of the Relative Position of Figure to Ground

*Its location is 30 miles north of the East-West Germany border.*

### 3.4 Representing the Spatial Relation

We use the OBJECTLOC-RELPOS-RELTO case frame, adapted from [1], for representing the relative location of one object to another. Since an entity, be it a city, border or region, may not be viewed as synonymous with its geographic location, (specified as a location, path or boundary, respectively), the network built to represent the relative position of figure to ground is constructed between the *values* of geographic properties.

For example, Figure 12 is the network built to represent the proposition that some located object's location value which is 12.694 degrees east longitude, 51.109 north latitude is *160 miles south* of the location value of the reference object which is 12.716 degrees east longitude, 51.316 degrees north latitude. The relative position is a structured individual which is denoted by a MEASURE-DIRECTION case frame that represents the distance and direction components. Distance, in the MEASURE slot, is also given as a structured individual with two components, the units used in the UNIT slot, and the number of them used in the VALUE slot.

The OBJECTLOC-RELPOS-RELTO case frame depicted in Figure 12 is expressing the relative position of the figure to a ground object with a location. If the ground is a region and hence possesses a boundary property, the RELTO arc is directed into the value of the boundary property. If the ground is a border with a path property, the RELTO arc is directed into the location property of the reference point chosen on the path. Hence, here the spatial relation is viewed as existing between an point on the path and the figure rather than the path itself and the figure. As discussed above, if the figure object is visually perceived as being so close to the ground that it appears next to the ground, the distance information given in the MEASURE slot is omitted.

## 4 Step Three: Natural Language Expression of the Spatial Relationship

The generation of natural language to express the spatial relationship represented by a node such as that in Figure 12 involves a straightforward traversal of the semantic node built to represent the spatial information. A generation grammar written in the Augmented Transition Network (ATN) formalism performs the traversal of the OBJECTLOC-RELPOS-RELTO case frame inserting words into an output string corresponding to sensory nodes at the ends of LEX arcs.



In addition to words, our extension to the generation grammar inserts LISP forms into the output string that when evaluated, produce visible gestures on the map. If the figure object is viewable, it highlights, labels and blinks the object's icon. We include the pointing/blinking gestures in the output string so that the gestures can be synchronized with the words as they are spoken:

The Erfurt air base is located here, < *blink air base* > 20 miles east of the East/West Germany border. < *arrow from border to air base* >

The grammar also briefly puts up the icon corresponding to ground object and draws an arrow from ground to figure. If the figure object is not viewable and the ground object is a region, the frame of the window in which the region is displayed is blinked.

## 5 Future Work

This work suggests five research activities, as referred to in this report.

- direction determination between two points
- direction determination from a path to a point
- preferential selection of ground objects from the focus list
- the representation of subregions and selection of associated landmarks appropriate to scale
- the mapping of the figure and ground objects to concepts that allow for the generation of acceptable natural language

We conclude with a brief description of the issues each is concerned with.

### 5.1 Direction Determination between Two Points

As previously noted, we suspect that determination of direction from one point to another with octants is not consistent with how direction is perceived. With respect to a map, the directions that appear to be acceptable for generating colloquial expressions are: *east, west, north, south, southeast, northeast, southwest, and northwest*. Visually however, the orientation of one point to another seems to be classified more often by one of the last four relations.

The possibility also exists that octants provide an acceptable division of space immediately around the ground point, but that as the distance between ground and figure becomes large relative to the display window, sensitivity to the last four directions increases. Data collected from subjects might help to determine the optimal division of angular space about a point for informal direction giving on a map.

### 5.2 Direction Determination from a Path to a Point

The problem here is perhaps twofold. As previously discussed, we feel that people are predisposed to certain directions based on geographic entities that can be characterized as paths. However we could not find a way to incorporate that kind of information into an algorithm for determining the direction from a border (represented as a path-like series of line segments) to a point (represented by coordinates given as latitude and longitude).

The second possible problem is related to the first. The computation of direction from a border currently involves computation of direction from a point chosen on the border as a reference point. Inasmuch as direction determination from a point to another point is a problem, so may direction determination from a border to point.

### 5.3 Preferential Selection of Ground Objects from the Focus List

As discussed, in selecting a ground object preference is given to a landmark that was previously used as the ground object and that is reasonably near the current object being located. Among the landmarks on the focus list that satisfy this criteria, the one closest to the figure is selected.

We suspect that this may be heavy handed. It appears to be the case that after a few initial exchanges, one or two landmarks become the dominant ground objects that all figure objects are located with respect to. This situation can be alleviated by enlarging a subregion that eliminates the dominant ground object, but does solve the problem.

However, this is a problem we feel has some quick and easy solutions. It may simply require changing the definition of what it means for a ground object to be "near" the figure object. We currently use half the window width, it may be that one-third or a quarter window width gives better results.

Another possibility is to restrict use of landmarks on the focus list to those of some certain minimum weight. Currently, we allow any landmark on the focus list that meets the criteria to be used, but it may be more cognitively pleasing if reuse of landmarks were restricted to within one or two exchanges. Their significance as landmarks may diminish more rapidly than they might actually fall off the focus list.

## 5.4 The Representation of Map Subregions

Typical of most map systems is the ability to pick an area of the map to enlarge. The subregion selected by the user is enlarged and displayed with greater refinement of existing features and addition of new ones. For example, in the CUBRICON system the user can ask to have a subregion enlarged and displayed in a different window. As an area is enlarged, cities that were previously represented by points become polygonal solids, rivers that were thin brown lines become bold blue ones that can be seen to run through cities. New features are also added. Smaller towns and highways that were not visible at a larger scale are represented at various levels of enlargement.

As the scale of the map changes, obviously so should the answer to a "Where is" query. If the addressee is looking at a map of New York State, the location of the University of Buffalo is well described as *near Buffalo*, a major city, or perhaps, in *western New York* depending on what is depicted. However, if the map displays the town of Amherst (the township where the campus is located), the response generated should clearly be in terms of local landmarks: *at the intersection of Maple and Sweet Home Roads* is one possibility.

Information on all the features that the map system is capable of displaying could be stored in the knowledge base. However, this would be a wasteful and cumbersome duplication of large volume of information. A better approach seems to be to build representations of subregions on the fly, as they are needed, and associate landmarks with them that are appropriate to scale and that are visually prominent.

A strategy for selecting landmarks appropriate to subregion scale is needed. This could also involve an imaging problem since a landmark, while present and appropriate to scale, might not be visually prominent in the context of other displayed features.

## 5.5 Mapping of Figure and Ground Objects to Geometric Descriptions

The locative expression used in part reflects how the object is perceived. With respect to the map context, as a feature is enlarged, it can visually undergo a transformation that can change how we talk about it. A city viewed as a point on a large scale map can become a boundary denoting city limits when enlarged, and as the image changes, so does how we talk about it. When viewed as a point: *at, near, east of*, etc. suggest themselves as candidate relationships the city is allowed to be apart of. When viewed as a boundary, *in* and *out of the city* become acceptable references to make. Clearly the choice of expressions for a spatial relationship is influenced by the perception of the objects involved.

Even if a precise representation of every spatial relationship were possible, its expression in natural language would still pose a problem. Herskovits suggests that for the purposes of expression, spatial relations are viewed as not existing between the objects themselves but between what she calls the *geometric descriptions* of objects. For example, the locative expression

*the village on the highway*

is intended to convey to the addressee that the highway runs through or near a village close by. The exact image evoked is different for each individual, however the natural language encoding of the idea is uniform: the village is conceptualized as a *point* and the highway as a *line*, and the point is *on* the line. The reverse process appears to take place for decoding or understanding the expression.

We need to address the issue of when and how this mapping should take place. As parties interested in knowledge representation, the issue of when to do the mapping is of particular interest. Two possibilities come to mind:

1. Do the mapping at the time of encoding
2. Do the mapping at the time of decoding

The first approach requires that the spatial relationship in the knowledge base be built between the appropriate geometric descriptions of the objects. Hence, in our example, the *on* relationship would exist between a point and line conceptualization of the village and highway respectively. Knowledge systems that support intensional representation, such as SNePS, are capable of implementing this strategy easily. The second approach requires less work in terms of representing the spatial relation, and is the better strategy to use in an extensional knowledge representation system. In the knowledge base, the spatial relation could be represented as existing between the canonical representations of the objects themselves.

The problem of how to do the mapping is discussed in detail by Herskovits. She applies a series of function compositions to an object to map it to a geometric description that is consistent with the *ideal meaning* of the preposition. However, the solution is not clean, she notes that these mappings are influenced by metonymic shifts and various contextual factors, many of which she attempts to explain with *near pragmatic principles*.

## **6 Conclusion**

Our work has resulted in the development of suitable representations for expressing spatial relations and has been implemented as a GATN grammar extension to the CUBRICON system that is capable of choosing a reference point and expressing the location of a geographic entity using the selected reference point in suitable natural language. The algorithm for landmark selection not only considers what is visually "near by" and prominent, it also takes into consideration what landmarks have been used recently to orient the user. In addition, the system expresses the observed relationship between the ground and figure object in natural language enhanced with graphic gestures. These are synchronized with the language to impart as much information as possible to the addressee with multi-modality responses. Finally, this work has clarified further research goals related to supplying visual information enhanced with appropriate verbal information.

## References

- [1] James Geller. *A Knowledge Representation Theory for Natural Language Graphics*. PhD thesis, State University of New York at Buffalo, 1988.
- [2] Annette Herskovits. Semantics and pragmatics of locative expressions. *Cognitive Science*, 9:341378, 1985.
- [3] Jeannette G. Neal and Stuart C. Shapiro. Intelligent multimedia interface technology. In Joseph W. Sullivan and Sherman W. Tyler, editors, *Proceedings of the Workshop on Architectures for Intelligent Interfaces: Elements and Prototypes*, pages 6991, 1988.
- [4] Stuart C. Shapiro. Generalized augmented transition network grammars for generation from semantic networks. *American Journal of Computational Linguistics*, 8:1225, 1982.
- [5] Stuart C. Shapiro and William J. Rapaport. Sneps considered as a fully intensional propositional semantic network. In Nick Cercone and Gordon McCalla, editors, *The Knowledge Frontier, Essays in the Representation of Knowledge*, pages 263315. SpringerVerlag, New York, 1987.
- [6] C. L. Sidner. Focusing in the comprehension of definite anaphora. In M. Brady and R. C. Berwick, editors, *Computational Models of Discourse*. MIT Press, Cambridge, 1983.