Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing

David M. Mark[1], David Comas[2], Max J. Egenhofer[3],
Scott M. Freundschuh[4], Michael D. Gould[5], Joan Nunes[6]

Abstract. Human-subjects testing based on a sound formal model is outlined as an effective way to evaluate and refine computational models of spatial relations. The cognitive response of a subject (person) to a real-world situation depends on the characteristic of the situation and the characteristics of the subject. Spatial entities can be characterized by topological relations, metric, scale (scope), kinds of phenomena (semantics), motion (time), and (for maps and displays) graphic presentation. Subjects' responses may vary according to their natural languages, their cultures, and individual characteristics such as gender, age, and handedness. Given this conceptual background, seven human-subjects testing protocols are introduced. Each test is described, and results of applications of these tests are summarized where possible. The tests are compared with their requirements, and the different aspects of human spatial cognition that they might test. Lastly, a program for applying these tests and refining the formal models based on test results is presented.

KEYWORDS: Spatial relations, spatial language, spatial cognition, formal models, human-subjects testing, geographic information systems, GIS.

---

1 National Center for Geographic Information and Analysis, Department of Geography, State University of New York at Buffalo, Buffalo, NY 14261-0023, U.S.A. Email: dmark@geog.buffalo.edu.

2 GIS Centre, Universitat de Girona, Pl. Ferrater Mora, 1, 17071 Girona, Spain. Email: comas@skywalker.udg.es.

3 National Center for Geographic Information and Analysis, Department of Spatial Information Science and Engineering, and Department of Computer Science, University of Maine, Orono, ME 04469-5711, U.S.A. Email: max@mecan1.maine.edu.

4 Department of Geography, University of Minnesota-Duluth, Duluth, Minnesota 55812, U.S.A. Email: sfreunds@d.umn.edu.

5 Departamento de Geografía y Ordenación del Territorio, Universidad de Extremadura, Avda. Quijotes, s/n, E-10071 Cáceres, Spain. Email: ghhum02@sis.ucm.es.

6 Departament de Geografia, Edifici B, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. Email: ILGE5@cc.uab.es.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

1. Introduction

Development and enhancement of a general theory of spatial relations are critical to several fields. Computational models of spatial relations are essential for the design of geographic information systems (GISs). Spatial relations are used internally by GISs in various overlay procedures. They also may be provided to end-users of the GIS as spatial predicates to constrain database retrievals and logical selections. But a general theory of spatial relations is also needed in cognitive science, especially in cognitive and computational linguistics. Spatial language appears in a wide range of linguistic areas, and is also a source domain for metaphors that people use to structure and talk about more abstract domains. Also, reasoning about spatial relations is very important in robotics and in computational vision. And lastly, a general theory of spatial relations would be of central interest to geography itself, for what it reveals about the nature of geographic space and spatial relations.

There has been much progress toward theories of spatial relations, notably the works of Talmy (1983) and Herskovits (1986) in linguistics and cognitive science, of Cohn and his colleagues in artificial intelligence (Randell *et al*., 1992; Cohn *et al.*, 1994; Cohn and Gotts, 1995), and of Egenhofer and his colleagues in geographic information science (Egenhofer and Franzosa, 1991; Egenhofer and Herring, 1994; Egenhofer and Mark, 1995). However, all the above work, and other related work in linguistics, artificial intelligence, mathematics, and GIS has been based on the intuition of the researchers. A model is accepted and published if it is internally consistent, intuitive, and makes plausible predictions or distinctions about spatial relations.

Recently, some of us have felt the need to conduct human-subjects experiments to confirm these formal models of spatial relations, or in some cases, to refine them. The resulting research approach seems powerful. The formal model is used to define things that can be distinguished in objective descriptions of situations, and experimental stimuli are designed to span the full range of situations that the formal model can distinguish. Test results may show that all of those distinctions are actually made by subjects, or that only some are made, or that some are made more frequently than others. What is learned can be fed back into the further development of the formal models. Basically, we believe that a person's reaction to a situation can be predicted in part from a formal description of a situation, and in part from characteristics of the subject (Figure 1). Our goal is to build such a predictive model, or suite of models, for spatial relations, with emphasis on relations in space of a geographic scale, and for how the relations are described or expressed in natural language.

If this research project is successful, the results would provide the basis for computer programs that could solve such problems as these:

- take an objective description of the geometry and attributes of two or more spatial entities, and the general characteristics of some person who is expected to understand the sentence, and determine an appropriate natural-language sentence or phrase that describes the spatial relation between the entities.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

- take any sentence involving locatives, any objective description of a situation involving spatial relations, and any general class of person, and determine how well the sentence describes the situation.

The computational ability to do the above would in turn allow, for example, retrieval from a spatial database of all those examples that fit a given sentence up to some specified degree. And, in the grand scheme of this project, this will be possible for all the world's natural languages.

In this paper, we present in some detail an outline of the factors that may be needed to describe spatial entities and their spatial relationships, using formal models that have sufficient expressive power to encompass everything that we need for building the models described above. Then, we will outline some characteristics of human subjects and eventual computer users, that also will have to be modeled. And third, we present a portfolio of human-subjects tests that we are performing, to provide the evidence to confirm, refine, and perhaps extend the formal models of spatial relations already developed.

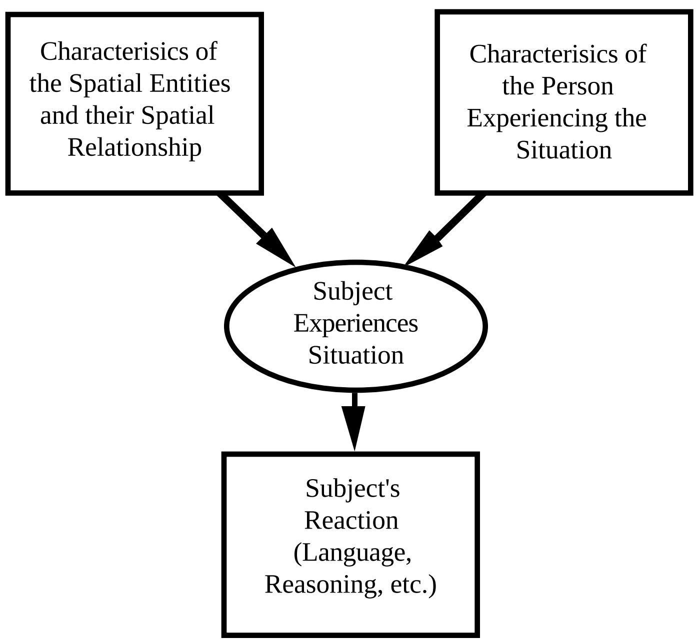


Figure 1: Diagram representing the framework for this research.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

2. Characterizing Spatial Entities and Their Relationships

2.1. Topology

The term topology refers to properties that are invariant under continuous transformations such as rotation, re-scaling, and stretching. Containment and coincidence are two of the more basic topological relations. In the research described in this paper, we define topological spatial relations as described by Egenhofer and Franzosa (1991) and Egenhofer and Herring (1994). Specifically, we use the 9-intersection model (Egenhofer and Herring, 1994), in which each entity is considered to have an interior, a boundary, and an exterior. In its simple form, the 9-intersection considers which of the three parts, for one entity intersect which parts of the other. Each of these 9 intersections can be either empty or non-empty. This would appear to lead to 512 distinct relations, but if each entity is simply connected, most combinations of part intersections are impossible. Egenhofer and Herring presented proofs on which combinations are impossible: the 9-intersection can distinguish in R^2 only eight spatial relations between regions in the plane, but 19 relations between an unbranched line and a region, and 33 between two unbranched lines.

The testing protocols described in this paper are not bound to the 9-intersection, but could be applied to stimuli chosen according to other models of topological differences. However, our thesis is that those aspects of spatial relations that can be differentiated using the 9-intersection topology are primary in defining spatial relations that make sense cognitively, and that these spatial relations form the basis of both natural-language processing and spatial query languages. The 9-intersection also can readily be extended by paying attention to the numbers of intersections, the dimensions of the intersections (e.g., intersection at a point or along a line), and the directions (orientations) of lines (Egenhofer and Franzosa, 1995).

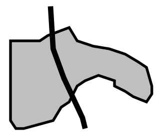
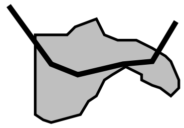
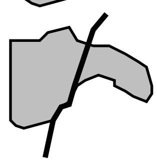
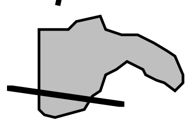
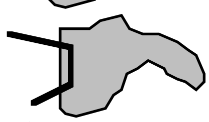
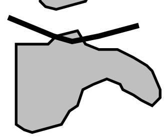
2.2. Metric

Whereas both theory and evidence suggest that spatial relations will be defined primarily by topology, we expect that metric differences also will play a role, and may dominate in some situations. Under metric we include the shapes and relative sizes of the entities, the sizes and angles of the intersections or of parts formed by the intersections, distances and directions between disjoint entities, relative and absolute orientations, and positions of one entity within another.

The shape of both entities involved in a spatial relationship can influence the relationship as judged by human subjects or even for purely computational purposes. More important may be the exact geometry of the spatial relation between the two features. For example, we have asked subjects the degree to which they agreed that in each case, "the road crosses the park," with a rating of 5 indicating strong agreement, and 1 strong disagreement. Ratings were then averaged across subjects and re-scaled to a 0 to 1 scale (Mark and Egenhofer, 1995), with a mean of 1.0 indicating unanimous strong agreement, and 0.0 unanimous strong disagreement. Six of the 60 stimuli that we have presented to about 120 English-speaking subjects were topologically-identical according to the 9-intersection model, but showed metric differences (Table 1). The three cases in which the road traversed the park between two opposing sides had mean agreement ratings of 0.88 and 0.91 that the road did "cross" the park, whereas for other stimuli, topologically identical to these, mean agreement ratings were as low as 0.28.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

Table 1: Mean Agreement that "The Road Crosses the Park" for Six Topologically-Identical Stimuli Over a Range of Geometries.

Stimulus	Agreement ("crosses")
	0.914
	0.886
	0.882
	0.781
	0.367
	0.285

The orientation of the Figure with respect to the Ground also may be important in defining spatial relations. The best example of "cross" may have a line running perpendicular to the long axis of a region that the line crosses, whereas a line that parallels the long axis may not fit the concept of "crosses" quite as well. Also, in testing based on diagrams, the orientation of the Figure and Ground with respect to the testing medium, be it a piece of paper or a computer screen may effect results, since a line that runs across the medium may have a stronger rating for "crosses" than one running up and down the page, other things being equal (Wagner, 1993). Geographic reference frames (north-south, east-west, etc.) may have similar effects, although this conjecture is completely untested.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

2.3. Scale/Scope

There is considerable evidence that people reason about small spaces such as table-tops, and the manipulable objects in them, differently than they reason about geographic spaces. Mark and Freundschuh (1995) recently reviewed the evidence that there are different kinds of cognitive spaces, based on the way people interact with them, but defined largely by scale (size, scope). Manipulable spaces are populated with three-dimensional entities that move without changing shape or other properties. Entities mostly have physical existence, and thus two things cannot occupy the same space. Geographic spaces, on the other hand, are mostly two-dimensional, and entities in them are viewed in a more abstract way, so that the same place can be occupied by, say, a province and a mountain at the same time.

Talmy (1983) has noted that almost all spatial relation terms in all languages apply across all scales. This is almost certainly true for the lexical elements, but the meanings of these words or case structures or whatever may change, both formally and informally, with scale or scope.

Testing using maps is somewhat problematic in this regard, since the map itself and its elements are clearly manipulable, but the map normally represents a geographic space and geographic entities. If there is a difference in the details of the meanings of spatial relations with scale, then results of map-based tests might relate to either space, or a mixture. In the long term, the research program must include experiments in which stimuli with identical geometries and topological relations, but where some subjects get them labeled as maps, and others get them labeled as diagrams or table-top-objects.

2.4. Kinds of Features and Phenomena

We believe that the nature of the phenomena involved can have an influence on how the spatial relations are perceived and characterized. We do not have experimental evidence to support this conjecture, as all experiments performed so far have involved a "road" and a "park." Furthermore, our conjecture conflicts with a generalization presented by Talmy (1983), who claimed that spatial relations in natural language tend to be both magnitude (size) neutral and material neutral. An example, based on introspection, was presented by Mark and Egenhofer (1994a):

> [T]he following pair of sentences would seem to mean about the same:
>
> The road goes through the park.
> The road goes across the park.
>
> The near-identity of "through" and "across" also seems to hold if we are talking about a man walking though or across a field. There may be a slight shade of difference in the meanings of the next pair of sentences, but again, the meanings seem very similar:
>
> The highway goes through Boston.
> The highway goes across Boston.
>
> But now consider this pair of sentences:
>
> When we went to Seattle, we drove through Canada.
> When we went to Seattle, we drove across Canada.
>
> The first sentence would be true if any part of the route is in Canada. It would, for example, be reasonable to say that if we had driven from Boston to Buffalo, then cut through southern Ontario,

> and then driven on U.S. roads from Detroit to Seattle. However, the second sentence implies that most of the route was in Canada, that is, from one "side" of Canada to the other.

We also expect that the nature of the boundaries of regions may play a role in the grouping or similarity of spatial relations. Recently, interest has focused on the fact that many geographic entities have indistinct boundaries (Frank, 1995; Burrough and Frank, 1995). Experiments performed so far, with a "road" and a "park," involved entities with crisp or distinct boundaries, and data suggest that subjects often made distinctions based on exact details of whether, say, the end of the road (line) fell precisely on the boundary of the park (region). Although we have not yet tested other phenomena, we suspect that if boundaries are thought by the subjects to be uncertain or imprecise, then the exact relations between two boundaries may not be of high salience.

2.5. Motion (Dynamic/Static)

Some languages make a major distinction between static and dynamic spatial relations. For example, Korean typically expresses spatial relation through trajectory verbs, rather than through static locative particles (Bowerman, 1993). Pilot data in this research have shown small differences in subjects' responses to "the road crosses the park" and "the road goes across the park," with mean agreement by subjects being weaker for the second sentence in the case of 48 of 60 stimuli, although differences were statistically significant (at the 5 percent level) for only three of the 60 stimuli (exactly five percent of them). Such a grammatical distinction may make a more substantial difference in some languages.

2.6. Presentation

It is fairly common to perform tests using maps or map-like drawings as stimuli, and then to draw conclusions about geographic terminology from the results. This may work, reasonably well, most of the time, but it must be remembered that experiments conducted in this way really only reveal things about how people talk about maps and the spatial relations shown on them. In fact, the situation may be worse than that, since most tests of spatial relations, spatial knowledge acquisition, or navigation have used as stimuli schematic drawings that do not look much like cartographic maps. It seems possible that subjects may lose track of the map metaphor during the task, and end up making judgments about points, lines, and polygons. For example, Mark and Egenhofer (1994b) showed cards with a closed polygon (shaded) and a solid line on each, saying that these represented a road and a state park. The subjects grouped the cards and then described the spatial relation in each group. In the 97 descriptions composed by 11 English-speaking subjects, 93 did use the terms "road" and "park" for the line and region, respectively. However, in the remaining six descriptions, coming from two of the subjects, the "road-park" (map) context broke down, and the subjects referred to the entities as the "line" or the "area." Thus, issues of the medium of presentation, the style or graphic design of the stimuli, and the presence or absence of context or base-map information, all are important variables in research designs in this area.

3. Characteristics of Subjects that may Influence Their Reactions to Spatial Relationships

As noted in the introduction to this paper, objective formal description of a spatial relation is only one side of the equation. A subject's reaction to a situation, and by inference, reactions of people to spatial relations in everyday life, also is influenced by the characteristics of the person. At present, we divide the individual characteristics into language spoken, culture, and individual characteristics.

3.1. Language Spoken

Language spoken has an obvious surface effect, but whether language has an effect on spatial reasoning and judgments is less certain and more controversial (see Montello, 1995). Benjamin Lee Whorf and Edward Sapir proposed that people think in language, and that speakers of different languages think and reason differently (Whorf, 1940; for a recent discussion in a geographic context, see Mark, 1993). This conjecture, known now as the Sapir-Whorf hypothesis is very controversial, because if such differences do exist, then it could be argued that speakers of one language are better equipped to think about certain topics than are speakers of some other language. Whereas there do seem to be differences in how languages categorize geographic features (Mark 1993), it is not at all clear whether such effects would extend to the definitions and groupings of spatial relations. Much data, using a variety of experimental protocols, and for many languages from as many language families as possible, will be needed before such effects can be confirmed or laid to rest.

3.2. Culture

> "It is impossible to comprehend a culture without taking into account its language, probably its single most important element. It is also impossible to completely understand a language independent of its cultural context. … But despite the many ways in which culture and language influence each other, their integration is not absolute. Each has many properties uniquely its own that are not directly, or even indirectly, influenced by the other." (Howard, 1989, p. 77)

Oft-cited examples of cultural influences on language are the "Eskimo words for snow" idea, in which the Inuit are claimed to have many words for kinds of snow, and no single word for everything that an English-speaker would consider to be snow. Whereas the facts regarding snow and the language of the Inuit seem unclear, the idea that some cultures have more detailed vocabularies than others seems to be true at least to some degree. The skiing "sub-culture" or "speech community" within North American English also has many words for snow, some of which might not be familiar to some non-skiing native speakers of North American English from snow-free localities. Whether there is any analogous variation in words for spatial relations is not known at this time, but is something to be looked for in interpretations of test results. We must be careful not to generalize about English speakers from undergraduate student subjects in Buffalo, New York, or about Spanish from office workers in San José, Costa Rica.

3.3. Individual Characteristics

Sex (gender) appears to be correlated with both verbal and spatial abilities (see Halpern, 1986, for an excellent summary). Since the tasks outlined in this paper involve judgments of spatial relations, or sentences, or both, we expect that both verbal and spatial abilities may influence results, and thus a gender effect is possible. We have asked all subjects in all the experiments so far to report their age, sex, and native (first) language spoken, although we do not expect significant sex-related differences in the results of these particular experiments. Direct testing of subjects' spatial and verbal abilities would be desirable, to see if that correlates with any aspects of their responses. Since Halpern (1986) noted significant interactions between cognitive abilities, sex, and handedness, it might also be worthwhile to record the handedness of subjects, although we have not done that so far.

4. Human-subjects Experiments for Assessing Spatial Relations

We have designed seven human-subjects protocols for exploring, evaluating, or refining computational models of spatial relations in natural language. All of these have been approved by the Human Subjects Review Board at the State University of New York at Buffalo. We have decided to use a road and a park as the entities in the spatial relationship for all of the testing in the early phase of our research, although eventually will extend the research to other kinds of geographic features. As illustrated in Figure 2, four tasks require road-park drawings, and two tasks produce such drawings. Also, two tasks require road-park spatial relation sentences, and three others produce them. We have developed a standard set of 60 road-park drawings for use with the verbal description (4), grouping (5) and agreement (7) tasks; these have been printed in an appendix to a technical report (Egenhofer *et al.*, 1994), and can be made available in electronic form to interested researchers.

4.1. Verbal Examples Task

In the Verbal Examples Task, subjects are asked to list as many different road-park sentences as they can, in their native language. Sentence frequency will be counted by subject, and sentence sequences can be studied to see whether subjects list approximately synonymous sentences consecutively, or use some other recall strategy.

4.2. Graphic Examples Task

This task is very similar to the previous one, except that subjects are asked to draw examples of as many different spatial relations as they can. Results will be classified topologically according to the distinctions of the 9-intersection model, and metric properties of the examples might also be characterized.

4.3. Graphic Comparison Task

This protocol presents road-park drawings three at a time, and asks subjects to judge whether the middle example of the three is more similar to the one to the left, or to the one to the right. This is a "forced choice" task, and we expect results to indicate the relative importance of topology and metric. Since both this and the previous task do not involve language explicitly, it will be very interesting to see whether there are any systematic differences in responses for speakers of different languages. The stimuli that we will use first involve relations between rectangles and straight lines.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

Some subjects will be judging geometric figures with no physical interpretation, whereas others will be told that the line represents a road and the rectangle represents a park. Any difference between those two conditions would tentatively be attributed to the semantics of the phenomena involved.

4.4. Verbal Description Task

The Verbal Description Task asks subjects to write a sentence that describes the spatial relation shown in each of the 60 diagrams mentioned above. In future, we may design other sets of diagrams.

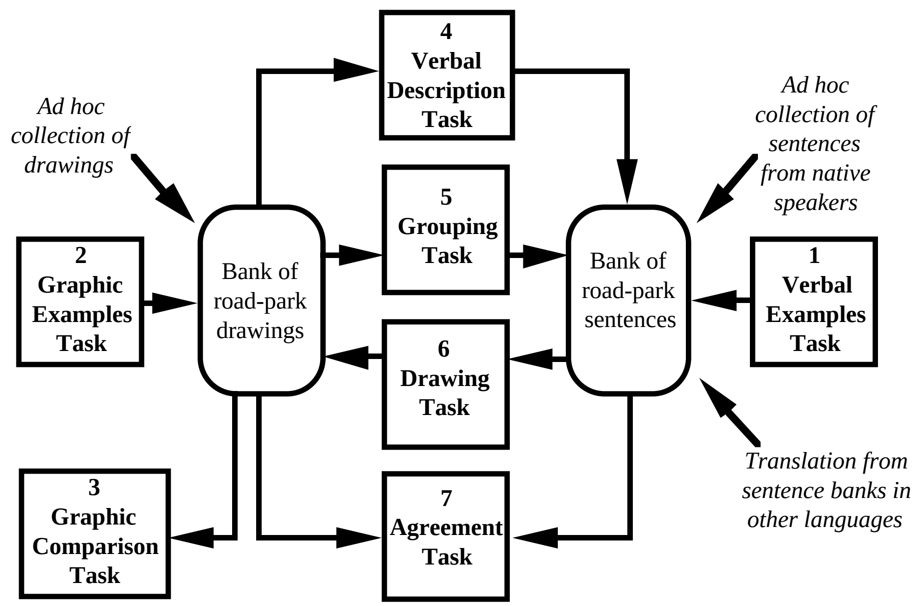


Figure 2: Information requirements and products from the seven experimental protocols defined thus far.

4.5. Grouping Task

In the Grouping Task, the 60 drawings are printed on individual cards, and shown to subjects, who are asked to arrange the drawings into groups such that a single natural-language phrase in their native language can describe each drawing in the group. To the maximum extent possible, the instructions are given and responses recorded in the native language of the subjects. In English, the instructions are:

> Here are 60 different sketches of a road and a State Park. Please arrange the sketches into several groups, such that you would use the same verbal description for the spatial relationship between the road and the Park in each case.

When the subject completes the task, the experimenter records the groups, and elicits a descriptive phrase for the spatial relation in each group. Lastly, the subject is asked to pick a "best example" from the stimuli in each group.

4.6. Drawing Task

The Drawing Task presents subjects with outlines of a park, with a sentence under each drawing. In English, the instructions are:

> On each of the following diagrams, the shaded polygon represents a state park. Please draw a line on each diagram to represent a road that has the spatial relationship to the park that is described by the sentence under the diagram. Try to draw a road that makes the diagram a "best example" of the relationship described by the sentence. If you think two sentences indicate the same spatial relation, you can draw the road in the same place to exemplify each.

Responses are categorized to determine which of the 19 possible spatial relations has been drawn. We also will study the geometry of the examples produced, noting whether arrows were added to "direct" the roads, etc.

4.7. Agreement Task

In the Agreement Task, subjects are presented with sentences, each of which describes a spatial relation between a "road" and a "park." The first page includes the following instructions:

> Examine each of the 60 maps, and determine how well you think the sentence printed at the top of side 1 of your answer form fits the spatial (geographic) relationship between the thicker dark road and the park. Your judgment should be on the scale of (a) "strongly disagree" to (e) "strongly agree."

This instruction page is followed by 8 pages, and each but the last of these pages had 8 road-park diagrams.

4.8 Comparison of the Tasks

The tasks described in sections 4.1 to 4.7 clearly have the potential to reveal different aspects of how subjects think about spatial relations. A good example of the difference between tasks is the contrast between the responses to the Drawing Task and the Agreement Task for the sentence "the road crosses the park" (Figure 3). In a sense, the results agree, since the topological class with the highest agreement rating for the sentence from the Agreement Task also was the class drawn by 29 of 32 subjects (over 90%) as an example of the sentence. However, although other topological classes had ratings from the Agreement Task that were almost as high, only 3 subjects drew any of these other patterns. Agreement values span the range from 0 to 1, showing that the category represented by "the road crosses the park" has graded or perhaps fuzzy membership. But the Drawing Task misses this, since the subject produces a prototypical or best example.

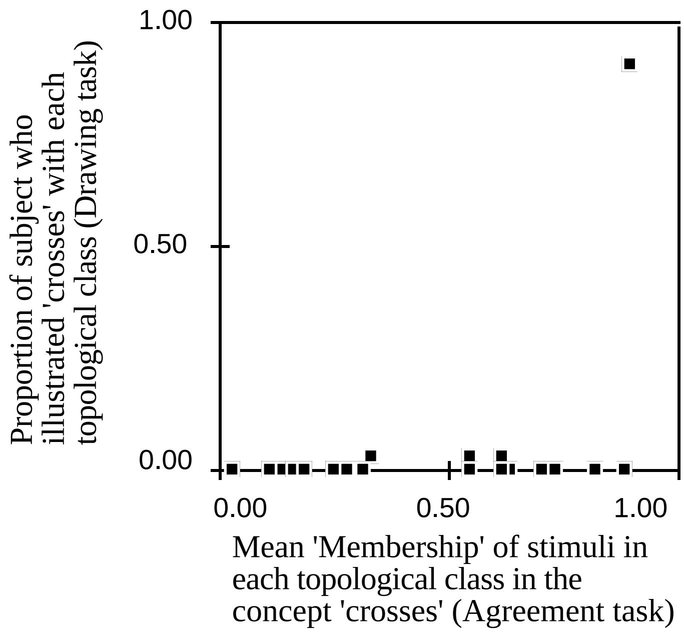


Figure 3: Plot of proportion of subjects who drew examples which each of the nineteen 9-intersection topologies as examples of "the road crosses the park" (vertical), plotted against mean agreement that "the road crosses the park" for examples of each of those topological classes.

4.9 Progress and Preliminary Results

Of the seven experimental protocols described in this paper, three have already been applied to groups of subjects.

The Grouping (5) Task was the first one that we used in this research program, asking 12 native speakers of English, 12 native speakers of Chinese, and four others, to group our first forty road-park stimuli into groups, such that every example within each group could be described by a simple sentence in their native language (Mark and Egenhofer, 1994b). The results showed a great deal of individual differences. Data derived from the results also were used to evaluate probable information loss that would result if the 9-intersection model were simplified to ignore certain topological distinctions (Egenhofer and Mark, 1995). Clearly, data from more subjects, and especially from more languages, are needed.

The Drawing (6) Task has been used to compare prototypical meanings of 64 sentences in English and 43 sentences in Spanish, with 32 and 19 subjects, respectively (Mark and Egenhofer, 1995). It seems that this task would be a useful early experiment in the exploration of a language new to these experimental procedures. The geometry of the drawings provides a rich set of data for exploring geometric factors in prototypical meanings.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

So far, the Agreement (7) Task has been applied to several sentences in English, Norwegian, and Spanish. Mark and Egenhofer (1994a, 1994b) reported results for four sentences in the English language, and four others have been tested subsequently. These are listed in Table 2. Abrahamson (1994) reported results for the Agreement Task for five sentences in Norwegian, and also we have tested one sentence in Spanish.

Table 2: English-language Sentences tested Using the Agreement Task, as of June 1995

the road crosses the park
the road goes across the park
the road goes through the park
the road enters the park
the road goes into the park
the road goes along the park
the road goes by the park
the road goes to the park

5. Summary and Further Work

In this paper, we have described a battery of human-subjects testing protocols that we feel will produce valuable insights into how people think about spatial relations and how they express these relations in natural language. Subject responses are expected to be associated with objective descriptions of the spatial entities and relationships, and on subject characteristics. Experimental design is based on the assumption that topology is primary in defining spatial relations, and that a particular formal model of topology, the 9-intersection, highlights the same distinctions that people make when judging and talking about spatial relations. Other aspects that may influence subjects' responses include metric, scale (scope), kinds of phenomena (semantics), motion (time), and graphic presentation. In the experiments conducted thus far, most of these have been held constant, but future work will involve systematic variation of test stimuli along these and perhaps other dimensions. Subjects' responses are expected to be influenced by native language, culture, and individual characteristics such as gender, age, and handedness.

Experiments that we have performed so far have been based on a line representing a road and a region representing a state park. Our plans for further testing employ a strategy that works outward from a core provided by the road-park situation in English. One dimension is to extend the tests to other sentences in English. Another is to test similar sentences in as many other languages as possible. And a third dimension is to change the phenomena involved from a road and a park to other geographic entity types. In each dimension, results can be compared back to the road-park-English core sentences. We will try to avoid the temptation to jump ahead into

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.

completely novel combinations of language, spatial relation, and phenomena, but instead attempt to build in an orderly fashion from the foundation provided by the formal model of topology and the tests performed thus far.

6. Acknowledgments

This paper is part of a collaborative research project funded in part by the Scientific and Environmental Affairs division of the North Atlantic Treaty Organization (NATO). The paper also is a part of Research Initiative 10, "Spatio-Temporal Reasoning in GIS," of the U.S. National Center for Geographic Information and Analysis (NCGIA), supported by the National Science Foundation (SBE-88-10917); support by NSF is gratefully acknowledged. Max Egenhofer's research is also supported by NSF under grant IRI-9309230, and by grants from Intergraph Corporation, Space Imaging Inc., and Environmental Systems Research Institute. Ann Deakin assisted with experiments at Buffalo and provided comments on the manuscript, and Jan Terje Bjorke coordinated the experiments with Norwegian.

7. References

Abrahamson, F., 1994. Uklarhet og Noyaktighet i Geografiske Informasjonssystemer (Fuzziness and Accuracy in Geographical Information Systems), Unpublished Master's thesis, Division of Surveying and Mapping, University of Trondheim (Norway) and Norwegian Institute of Technology.

Bowerman, M., 1991. The Origins of Children's Spatial Semantic Categories: Cognitive vs. Linguistic Determinants. In J. J. Gumperz and S. C. Levinson, Editors, *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Burrough, P., and Frank, A. U., (Editors) 1995. *Geographic Objects with Indeterminate Boundaries*. London, Taylor and Francis, in press.

Cohn, A. G., and Gotts, N. M., 1995. The "Egg-Yolk" Representation of Regions with Indeterminate Boundaries. In P. Burrough and A. U. Frank, Editors, *Geographic Objects with Indeterminate Boundaries*. London, Taylor and Francis, in press.

Cohn, A. G., Randell, D. A., and Cui, Z., 1994. Taxonomies of Logically Defined Qualitative Spatial Relations. In N. Guarino and R. Poli, Editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*. Kluwer, in press.

Egenhofer, M., and Franzosa, R., 1991. Point-Set Topological Spatial Relations. *International Journal of Geographical Information Systems* 5(2): 161-174.

Egenhofer, M., and Franzosa, R., 1995. On the Equivalence of Topological Relations. *International Journal of Geographical Information Systems* 9(2): 133-152.

Egenhofer, M., and Herring, J., 1994. Categorizing Topological Spatial Relations Between Point, Line, and Area Objects. In M. J. Egenhofer, D. M. Mark, and J.

R. Herring, 1994. *The 9-Intersection: Formalism and its Use For Natural-Language Spatial Predicates*. Santa Barbara, CA: National Center for Geographic Information and Analysis, Report 94-1.

Egenhofer, M. J., and Mark, D. M., 1995. Modeling Conceptual Neighborhoods of Topological Relations. *International Journal of Geographical Information Systems* , in press.

Egenhofer, M. J., Mark, D. M., and Herring, J. R. 1994. *The 9-Intersection: Formalism and its Use For Natural-Language Spatial Predicates*. Santa Barbara, CA: National Center for Geographic Information and Analysis, Report 94-1.

Frank, A. U., 1995. The Prevalence of Objects with Sharp Boundaries in GIS. In P. Burrough and A. U. Frank, Editors, *Geographic Objects with Indeterminate Boundaries*. London, Taylor and Francis, in press.

Halpern, D. F., 1986. *Sex Differences in Cognitive Abilities*. Hillsdale, N. J.: Erlbaum Associates.

Herskovits, A., 1986. *Language and Spatial Cognition: A Interdisciplinary Study of the Prepositions in English*. Cambridge, England: Cambridge University Press.

Howard, M. C., 1989. *Contemporary Cultural Anthropology* (Third Edition). Harper Collins Publishers.

Mark, D. M., 1993. Toward a Theoretical Framework for Geographic Entity Types. In A. U. Frank and I. Campari, Editors, *Spatial Information Theory: A Theoretical Basis for GIS*, Berlin: Springer-Verlag, *Lecture Notes in Computer Sciences* Vol. 716, pp. 270-283.

Mark, D. M., and Egenhofer, M. J., 1994a. Calibrating the Meanings of Spatial Predicates From Natural Language: Line-Region Relations. *Sixth International Symposium on Spatial Data Handling* 1994, Vol. 1, pp. 538-553.

Mark, D. M., and Egenhofer, M. J., 1994b. Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subjects Testing. *Cartography and Geographic Information Systems* 21(3), in press.

Mark, D. M., and Egenhofer, M. J., 1995. Topology of Prototypical Spatial Relations Between Lines and Regions in English and Spanish. *AutoCarto 12*, Charlotte, NC, March 1995, pp. 245-254.

Mark, D. M., and Freundschuh, S. M., 1995. Spatial Concepts and Cognitive Models for Geographic Information Use. In T. L. Nyerges, D. M. Mark, R. Laurini, and M. Egenhofer, Editors, *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*. Dordrecht: Kluwer Academic Publishers.

Montello, D., 1995. Cultural Differences in Spatial Cognition? *COSIT'95, Lecture Notes in Computer Science,Lecture Notes in Computer Science*, Berlin: Springer-Verlag, in press.

Randell, D. A., Cohn, A. G., and Cui, Z., 1992. A Spatial Logic Based on Regions and Connection. *3rd International Conference on Knowledge Representation and Reasoning*. San Francisco: Morgan Kaufmann, pp. 156-176.

Wagner, L., 1993. *The Mental Representation of the Meaning of Across*. Unpublished Undergraduate Thesis, Wellesley College.

Talmy, L., 1983. How Language Structures Space. In H. Pick and L. Acredolo, Editors, *Spatial Orientation: Theory, Research and Application.* Plenum Press.

Whorf, B. L., 1940. Science and Linguistics. *Technology Review (M.I.T.)* 42(6). reprinted in J. B. Carroll, Editor, 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, MA: The M.I.T. Press, pp. 207-219.

D. Mark, D. Comas, M. Egenhofer, S. Freundschuh, M. Gould, and J. Nunes, Evaluating and Refining Computational Models of Spatial Relations Through Cross-Linguistic Human-Subjects Testing, in: A. Frank and W. Kuhn (eds.), Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria, Lecture Notes in Computer Science, Vol. 988, Springer-Verlag, Berlin, pp. 553-568.